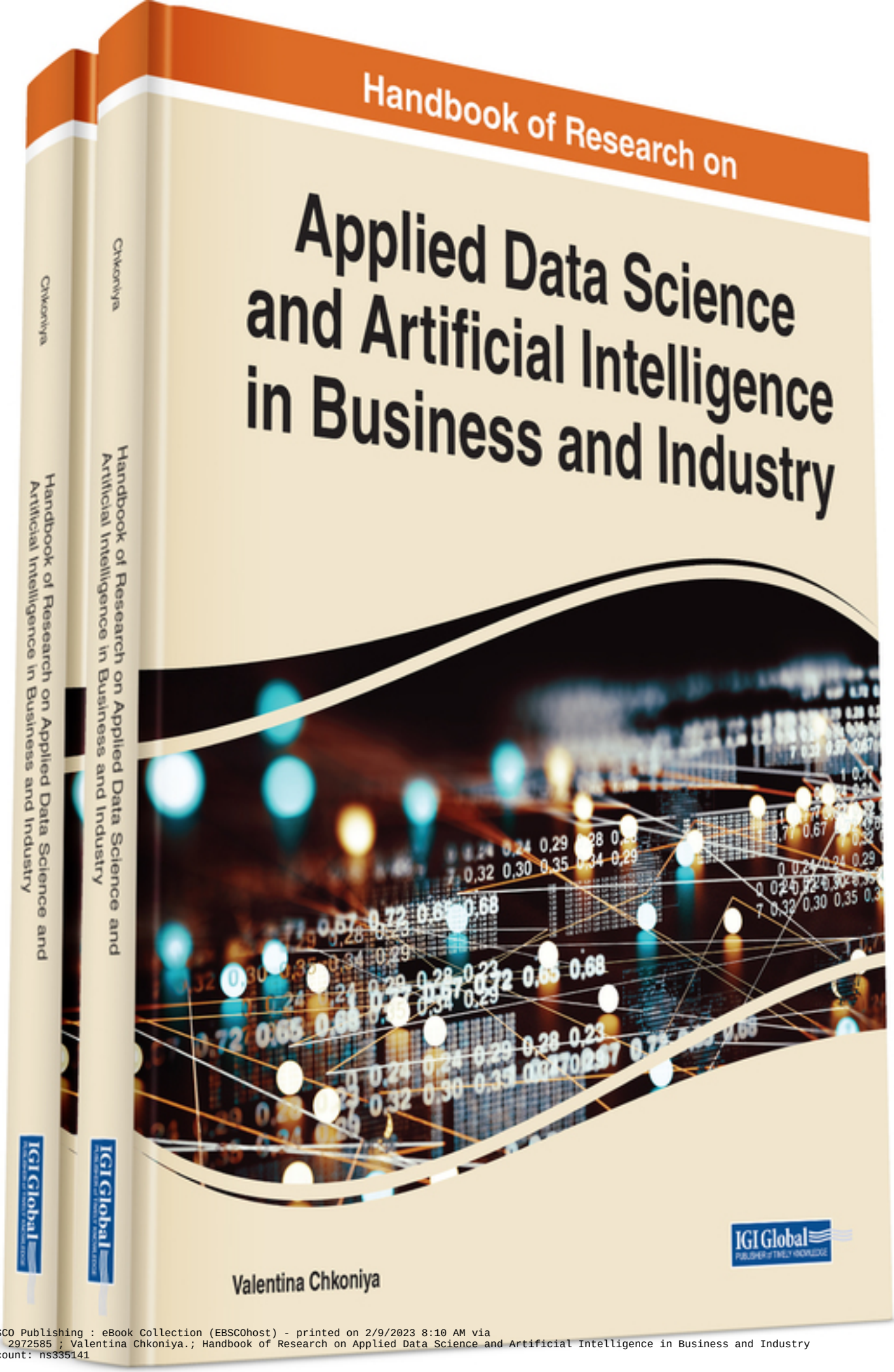


Copyright 2021. Engineering Science Reference. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/9/2023 8:10 AM via
AN: 2972585 ; Valentina Chkoniya.; Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry
Account: ns335141

Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry

Valentina Chkoniya
University of Aveiro, Portugal

A volume in the Advances in Business Information Systems and Analytics (ABISA) Book Series

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Published in the United States of America by
IGI Global
Engineering Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Copyright © 2021 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Names: Chkoniya, Valentina, 1977- editor.
Title: Handbook of research on applied data science and artificial intelligence in business and industry / Valentina Chkoniya, editor.
Description: Hershey : Engineering Science Reference, 2021. | Includes bibliographical references and index. | Summary: "This book provides an overview of the concepts, tools and techniques behind the increasing field of Data Science and Artificial Intelligence, applied to business and industries"-- Provided by publisher.
Identifiers: LCCN 2020058395 (print) | LCCN 2020058396 (ebook) | ISBN 9781799869856 (hardcover) | ISBN 9781799869863 (ebook)
Subjects: LCSH: Business--Technological innovations. | Databases. | Artificial intelligence.
Classification: LCC HD45 .H2934 2021 (print) | LCC HD45 (ebook) | DDC 338.00285/57--dc23
LC record available at https://lccn.loc.gov/2020058395
LC ebook record available at https://lccn.loc.gov/2020058396

This book is published in the IGI Global book series Advances in Business Information Systems and Analytics (ABISA) (ISSN: 2327-3275; eISSN: 2327-3283)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: eresources@igi-global.com.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Advances in Business Information Systems and Analytics (ABISA) Book Series

Madjid Tavana
La Salle University, USA

ISSN:2327-3275
EISSN:2327-3283

Mission

The successful development and management of information systems and business analytics is crucial to the success of an organization. New technological developments and methods for data analysis have allowed organizations to not only improve their processes and allow for greater productivity, but have also provided businesses with a venue through which to cut costs, plan for the future, and maintain competitive advantage in the information age.

The **Advances in Business Information Systems and Analytics (ABISA) Book Series** aims to present diverse and timely research in the development, deployment, and management of business information systems and business analytics for continued organizational development and improved business value.

Coverage

- Data Governance
- Business Models
- Business Intelligence
- Business Systems Engineering
- Legal information systems
- Big Data
- Information Logistics
- Business Information Security
- Performance Metrics
- Forecasting

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: http://www.igi-global.com/publish/.

The Advances in Business Information Systems and Analytics (ABISA) Book Series (ISSN 2327-3275) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, www.igi-global.com. This series is composed of titles available for purchase individually; each title is edited to be contextually exclusive from any other title within the series. For pricing and ordering information please visit http://www.igi-global.com/book-series/advances-business-information-systems-analytics/37155. Postmaster: Send all address changes to above address. Copyright © 2021 IGI Global. All rights, including translation in other languages reserved by the publisher. No part of this series may be reproduced or used in any form or by any means – graphics, electronic, or mechanical, including photocopying, recording, taping, or information and retrieval systems – without written permission from the publisher, except for non commercial, educational use, including classroom teaching purposes. The views expressed in this series are those of the authors, but not necessarily of IGI Global.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Titles in this Series

For a list of additional titles in this series, please visit: www.igi-global.com/book-series/advances-business-information-systems-analytics/37155

Business Applications in Social Media Analytics
Himani Bansal (Jaypee University, Solan, India) and Gulshan Shrivastava (National Institute of Technology, Patna, India)
Business Science Reference • © 2021 • 330pp • H/C (ISBN: 9781799850465) • US $195.00

Using Strategy Analytics to Measure Corporate Performance and Business Value Creation
Sandeep Kautish (Lord Buddha Education Foundation, Nepal)
Business Science Reference • © 2021 • 315pp • H/C (ISBN: 9781799877165) • US $225.00

Adapting and Mitigating Environmental, Social, and Governance Risk in Business
Magdalena Ziolo (University of Szczecin, Poland)
Business Science Reference • © 2021 • 313pp • H/C (ISBN: 9781799867883) • US $195.00

Innovative and Agile Contracting for Digital Transformation and Industry 4.0
Mohammad Ali Shalan (Aqarat Real Estate Development Company, Saudi Arabia) and Mohammed Ayedh Algarni (Information and Documents Center, Institute of Public Administration, Saudi Arabia)
Business Science Reference • © 2021 • 415pp • H/C (ISBN: 9781799845010) • US $225.00

Opportunities and Strategic Use of Agribusiness Information Systems
Ferdinand Ndifor Che (W3-Research, USA & APPC Research, Australia) Kenneth David Strang (W3-Research, USA & APPC Research, Australia) and Narasimha Rao Vajjhala (University of New York Tirana, Albania)
Business Science Reference • © 2021 • 333pp • H/C (ISBN: 9781799848493) • US $195.00

Integration Challenges for Analytics, Business Intelligence, and Data Mining
Ana Azevedo (CEOS:PP, ISCAP, Polytechnic of Porto, Portugal) and Manuel Filipe Santos (Algoritmi Centre, University of Minho, Guimarães, Portugal)
Engineering Science Reference • © 2021 • 250pp • H/C (ISBN: 9781799857815) • US $225.00

Managing Business in the Civil Construction Sector Through Information Communication Technologies
Bithal Das Mundhra (Simplex Infrastructures Ltd, India) and Rajesh Bose (Simplex Infrastructures Ltd, India)
Business Science Reference • © 2021 • 254pp • H/C (ISBN: 9781799852919) • US $195.00

701 East Chocolate Avenue, Hershey, PA 17033, USA
Tel: 717-533-8845 x100 • Fax: 717-533-8661
E-Mail: cust@igi-global.com • www.igi-global.com

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Editorial Advisory Board

Artur Alves, *Marloconsult, Portugal*
Elodie Attié, *Capgemini Engineering, France*
Belem Barbosa, *University of Porto, Portugal*
Paulo Batista, *University of Aveiro, Portugal*
Arnab Bhattacharjee, *Heriot-Watt University, UK*
Paata Bukhrashvili, *Ilia State University, Georgia*
Yezelia Danira Caceres Cabana, *Universidad Nacional de San Agustín de Arequipa, Peru*
José Luís Cacho, *Port of Sines, Portugal*
Gianluigi De Pascale, *University of Foggia, Italy*
Alice Duarte, *University of Porto, Portugal*
Francesco Facchini, *Polytechnic University of Bari, Italy*
Elizaveta Fainshtein, *National Research University Higher School of Economics, Russia*
Marcos Ferreira Santos, *Universidad de La Sabana, Colombia*
Sandra Filipe, *University of Aveiro, Portugal*
Adelaide Freitas, *University of Aveiro, Portugal*
Francisco Haro-Navejas, *Universidad de Colima, Mexico*
Marcel Kyas, *Reykjavik University, Iceland*
Joao Lourenço Marques, *University of Aveiro, Portugal*
Mara Madaleno, *University of Aveiro, Portugal*
Luís Jorge Gonçalves, *University of Aveiro, Portugal*
Raquel Meneses, *University of Porto, Portugal*
Jorge Mora Fernández, *University of California, San Diego, USA*
José Moreira, *University of Aveiro, Portugal*
Elissa Moses, *Columbia University, USA & BrainGroup Global, USA*
Ana Belén Nieto Librero, *University of Salamanca, Spain*
Helena Nobre, *University of Aveiro, Portugal*
Rafał Ohme, *WSB University in Torun, Poland*
Nino Okrostsvaridze, *Ilia State University, Georgia*
Roberto Rana, *University of Foggia, Italy*
Dorota Reykowska, *NEUROHM, Poland*
Eugénio M. Rocha, *University of Aveiro, Portugal*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nandini Sen, *Heriot Watt University, UK*
Dora Simões, *University of Aveiro, Portugal*
Adalberto Tokarski, *National Waterway Transportation Agency, Brazil*
Ana Tomé, *University of Aveiro, Portugal*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

List of Contributors

Aguilar del Carpio, David / *Universidad Nacional de San Agustín de Arequipa, Peru* 323
Albayrak, Özlem Karadağ / *Kafkas University, Turkey* 491
Altrad, Abdullah M. M. / *Al-Madinah International University, Malaysia* 431
Ashfaq, Muhammad / *Shenzhen University, China* 291
Attié, Elodie / *Capgemini Engineering, France* 619
Ayanso, Anteneh / *Brock University, Canada* 85
Bansal, Veena / *Indian Institute of Technology, Kanpur, India* 639
Baran, Tamer / *Pamukkale University, Turkey* 414
Barbosa, Belém / *University of Porto, Portugal* 366
Barinaga, Borja / *Universidad Francisco de Vitoria Madrid, Spain* 569
Batista, Maria Manuela Martins / *ENIDH, Portugal* 506
Bhatt, Chintan M. / *Charotar University of Science and Technology, India* 401
Bienvenido, Hector Puente / *Universidad Complutense de Madrid, Spain* 569
Bughin, Jacques R. J. / *Université libre de Bruxelles, Belgium* 126
Caceres Cabana, Yezelia Danira / *Universidad Nacional de San Agustín de Arequipa, Peru* 323
Cacho, José Luís / *Port of Sines, Portugal* 1
Chkoniya, Valentina / *University of Aveiro, Portugal* 1, 381, 506
Cincera, Michele / *Université libre de Bruxelles, Belgium* 126
Clark, Kimberly Rose / *Dartmouth College, USA & Bellwether Citizen Response, USA* 151
Cruz, Luís / *Delft University of Technology, The Netherlands* 23
Del Sarto, Nicola / *Scuola Superiore Sant'Anna, Italy* 108
El-Ebiary, Yousef A. Baker / *Universiti Sultan Zainal Abidin, Malaysia* 431
Fainshtein, Elizaveta / *National Research University Higher School of Economics, Russia* 223
Ferraz, Lisa / *University of Aveiro, Portugal* 366
Ferreira, Nelson / *Bosch Termotecnologia SA, Portugal* 263
Freitas, Adelaide / *University of Aveiro, Portugal* 55
Freitas, Ana Costa / *Sonae MC, Portugal* 168
Galindo, Purificación / *University of Salamanca, Spain* 55
Gallardo-Pérez, Julia / *University of Huelva, Spain* 304
Georgieva, Petia Georgieva / *University of Aveiro, Portugal* 263
Gonçalves, Fernando Cruz / *ENIDH, Portugal* 506
Gusmão, Pedro / *IEETA, University of Aveiro, Portugal* 241
Heck, Petra / *Fontys University of Applied Sciences, The Netherlands* 23
Hilles, Mohanad M. / *University College of Applied Science, Palestine* 431
Hilles, Shadi M. S. / *Istanbul Okan University, Turkey* 431

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Infante-Moro, Alfonso / *University of Huelva, Spain* 304
Infante-Moro, Juan C. / *University of Huelva, Spain* 304
Izquierdo-Yusta, Alicia / *University of Burgos, Spain* 192
Jacknis, Norman J. / *Columbia University, USA & Intelligent Community Forum, USA* 151
Jiménez-Hernández, Mónica / *Universidad de Colima, Mexico* 55
Jiménez-Zarco, Ana Isabel / *Open University of Catalonia (UOC), Spain* 192
Junior, Clony / *IEETA, University of Aveiro, Portugal* 241
Kyas, Marcel / *Reykjavik University, Iceland* 381
Lamest, Markus / *Trinity College Dublin, Ireland* 639
Le Bars, Solène / *Altran Lab, France* 619
Liban, Abdilahi Deria / *Hargeisa University, Somalia* 431
Lopes, Maria João / *Bosch Termotecnologia SA, Portugal & University of Aveiro, Portugal* 263
Loureiro, Sandra Maria Correia / *ISCTE, BRU, Instituto Universitário de Lisboa, Portugal* 291
Madaleno, Mara / *GOVCOPP, DEGEIT, Portugal & University of Aveiro, Portugal* 544
Marques, João Lourenço / *GOVCOPP, DCSPT, Portugal & University of Aveiro, Portugal* 544
Martínez-Ruiz, María Pilar / *University of Albacete, Spain* 192
Matias, Liliana Bernardino / *Sonae MC, Portugal* 168
Méndez-Aparicio, M Dolores / *Open University of Catalonia (UOC), Spain* 192
Mengesha, Niguissie / *Brock University, Canada* 85
Miranda, Filipe Fernandes / *Sonae MC, Portugal* 168
Mora-Fernandez, Jorge / *University of California, San Diego, USA* 569
Moreira, José / *DETI, University of Aveiro, Portugal* 241
Moses, Elissa / *Columbia University, USA & BrainGroup Global, USA* 151
Nobre, Helena / *University of Aveiro, Portugal* 366
Ohme, Rafał / *WSB University in Torun, Poland* 126
Patel, Parth / *Charotar University of Science and Technology, India* 401
Patel, Riddhi / *Charotar University of Science and Technology, India* 401
Pedersen, Jan Tore / *Marlo as., Norway* 381
Piccaluga, Andrea / *Scuola Superiore Sant'Anna, Italy* 108
Quenel, Ilhem / *Capgemini Engineering, France* 619
Rahevar, Mrugendrasinh L. / *Charotar University of Science and Technology, India* 401
Reykowska, Dorota / *NEUROHM, Poland* 126
Rivera Medina, Juan Mardonio / *Universidad Continental Perú, Peru* 323
Rocha, Eugénio M. / *University of Aveiro, Portugal* 263
Rodrigues, Mariana Oliveira Berga / *ISCTE, Instituto Universitário de Lisboa, Portugal* 291
Saxena, Deepak / *Birla Institute of Technology and Science, Pilani, India* 639
Schouten, Gerard / *Fontys University of Applied Sciences, The Netherlands* 23
Sen, Nandini / *Heriot Watt University, UK* 598
Serova, Elena / *National Research University Higher School of Economics, Russia* 223
Sığırcı, Özge / *Kırklareli University, Turkey* 342
Simões, Dora Maria / *University of Aveiro, Portugal* 522
Springer, Joshua D. / *Reykjavik University, Iceland* 381
Tejedor-Flores, Nathalia / *Universidad Tecnológica de Panamá, Panama* 55
Thomas, Elizabete / *National Waterway Transportation Agency, Brazil* 1
Tokarski, Adalberto / *National Waterway Transportation Agency, Brazil* 1
Tome, Ana Maria M. / *DETI, University of Aveiro, Portugal* 241

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Trad, Antoine / *IBISTM, France* 447
Tufail, Muhammad / *National University of Sciences and Technology (NUST), Pakistan* 544
Velásquez Chacón, Erika / *Universidad Católica San Pablo, Peru* 323
Vicente-Galindo, Purificación / *University of Salamanca, Spain* 55

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table of Contents

Preface xxvii

Acknowledgment xxxiv

Section 1
Strategy

Chapter 1
Port Dada Integration: Opportunities for Optimization and Value Creation 1
José Luís Cacho, Port of Sines, Portugal
Adalberto Tokarski, National Waterway Transportation Agency, Brazil
Elizabete Thomas, National Waterway Transportation Agency, Brazil
Valentina Chkoniya, University of Aveiro, Portugal

Chapter 2
A Software Engineering Perspective on Building Production-Ready Machine Learning Systems 23
Petra Heck, Fontys University of Applied Sciences, The Netherlands
Gerard Schouten, Fontys University of Applied Sciences, The Netherlands
Luís Cruz, Delft University of Technology, The Netherlands

Chapter 3
Multivariate Sustainability Profile of Global Fortune 500 Companies Using GRI-G4 Database 55
Mónica Jiménez-Hernández, Universidad de Colima, Mexico
Purificación Vicente-Galindo, University of Salamanca, Spain
Nathalia Tejedor-Flores, Universidad Tecnológica de Panamá, Panama
Adelaide Freitas, University of Aveiro, Portugal
Purificación Galindo, University of Salamanca, Spain

Chapter 4
Examining the Evolution of E-Government Development of Nations Through Machine Learning Techniques 85
Niguissie Mengesha, Brock University, Canada
Anteneh Ayanso, Brock University, Canada

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 5
Artificial Intelligence as Driver for SME Competitiveness 108
Nicola Del Sarto, Scuola Superiore Sant'Anna, Italy
Andrea Piccaluga, Scuola Superiore Sant'Anna, Italy

Chapter 6
Big Data Is Decision Science: The Case of COVID-19 Vaccination 126
Jacques R. J. Bughin, Université libre de Bruxelles, Belgium
Michele Cincera, Université libre de Bruxelles, Belgium
Dorota Reykowska, NEUROHM, Poland
Rafał Ohme, WSB University in Torun, Poland

Chapter 7
The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience..... 151
Elissa Moses, Columbia University, USA & BrainGroup Global, USA
Kimberly Rose Clark, Dartmouth College, USA & Bellwether Citizen Response, USA
Norman J. Jacknis, Columbia University, USA & Intelligent Community Forum, USA

Section 2
Action

Chapter 8
Retail Customer Native Baskets Creation: A Must-Do Into Measuring Price Indices 168
Liliana Bernardino Matias, Sonae MC, Portugal
Ana Costa Freitas, Sonae MC, Portugal
Filipe Fernandes Miranda, Sonae MC, Portugal

Chapter 9
Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence..... 192
M Dolores Méndez-Aparicio, Open University of Catalonia (UOC), Spain
María Pilar Martínez-Ruiz, University of Albacete, Spain
Alicia Izquierdo-Yusta, University of Burgos, Spain
Ana Isabel Jiménez-Zarco, Open University of Catalonia (UOC), Spain

Chapter 10
Using Intelligent Text Analysis of Online Reviews to Determine the Main Factors of Restaurant Value Propositions 223
Elizaveta Fainshtein, National Research University Higher School of Economics, Russia
Elena Serova, National Research University Higher School of Economics, Russia

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 11
Time Series Forecasting in Retail Sales Using LSTM and Prophet 241
Clony Junior, IEETA, University of Aveiro, Portugal
Pedro Gusmão, IEETA, University of Aveiro, Portugal
José Moreira, DETI, University of Aveiro, Portugal
Ana Maria M. Tome, DETI, University of Aveiro, Portugal

Chapter 12
General Model for Metrics Calculation and Behavior Prediction in the Manufacturing Industry: An Automated Machine Learning Approach 263
Maria João Lopes, Bosch Termotecnologia SA, Portugal & University of Aveiro, Portugal
Eugénio M. Rocha, University of Aveiro, Portugal
Petia Georgieva Georgieva, University of Aveiro, Portugal
Nelson Ferreira, Bosch Termotecnologia SA, Portugal

Chapter 13
AI Meaning and Applications in the Consumer Sector of Retailing, Hospitality, and Tourism 291
Sandra Maria Correia Loureiro, ISCTE, BRU, Instituto Universitário de Lisboa, Portugal
Muhammad Ashfaq, Shenzhen University, China
Mariana Oliveira Berga Rodrigues, ISCTE, Instituto Universitário de Lisboa, Portugal

Chapter 14
Key Factors in the Process of Acceptance and Implementation of Artificial Intelligence in the Hotel Sector 304
Alfonso Infante-Moro, University of Huelva, Spain
Juan C. Infante-Moro, University of Huelva, Spain
Julia Gallardo-Pérez, University of Huelva, Spain

Chapter 15
Peruvian Consumer Responses to the Use of Technologies in the Context of COVID-19 323
Yezelia Danira Caceres Cabana, Universidad Nacional de San Agustín de Arequipa, Peru
David Aguilar del Carpio, Universidad Nacional de San Agustín de Arequipa, Peru
Erika Velásquez Chacón, Universidad Católica San Pablo, Peru
Juan Mardonio Rivera Medina, Universidad Continental Perú, Peru

Chapter 16
Artificial Intelligence in Marketing: A Review of Consumer-AI Interactions 342
Özge Sığırcı, Kırklareli University, Turkey

Chapter 17
When Luxury Vinous-Concept Hotel Meets Premium Wine Brands: An Exploratory Study on Co-Branding 366
Lisa Ferraz, University of Aveiro, Portugal
Helena Nobre, University of Aveiro, Portugal
Belém Barbosa, University of Porto, Portugal

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 18
Data Analysis in the Shipping Industry: eShip Case Study – Problem Statement 381
Marcel Kyas, Reykjavik University, Iceland
Joshua D. Springer, Reykjavik University, Iceland
Jan Tore Pedersen, Marlo as., Norway
Valentina Chkoniya, University of Aveiro, Portugal

Chapter 19
Question Answering Chatbot Using Memory Networks 401
Riddhi Patel, Charotar University of Science and Technology, India
Parth Patel, Charotar University of Science and Technology, India
Chintan M. Bhatt, Charotar University of Science and Technology, India
Mrugendrasinh L. Rahevar, Charotar University of Science and Technology, India

Chapter 20
Comparison of Parametric and Non-Parametric Methods to Analyse the Data Gathered by a Likert-Type Scale 414
Tamer Baran, Pamukkale University, Turkey

Chapter 21
Latent Fingerprint Enhancement Based on EDTV Model..... 431
Shadi M. S. Hilles, Istanbul Okan University, Turkey
Abdilahi Deria Liban, Hargeisa University, Somalia
Abdullah M. M. Altrad, Al-Madinah International University, Malaysia
Yousef A. Baker El-Ebiary, Universiti Sultan Zainal Abidin, Malaysia
Mohanad M. Hilles, University College of Applied Science, Palestine

Chapter 22
The Business, Societal, and Enterprise Architecture Framework: An Artificial Intelligence-, Data Sciences-, and Big Data-Based Approach 447
Antoine Trad, IBISTM, France

Chapter 23
The Effect of Population, Public Investments, and Economic Growth on Regional Renewable Energy Consumption in Turkey: Dynamic Panel Data Analysis 491
Özlem Karadağ Albayrak, Kafkas University, Turkey

Section 3
Reflection

Chapter 24
Education for the Digital Industry: Opportunities and Challenges of Experience-Based Expertise and Open Innovation 506
Valentina Chkoniya, University of Aveiro, Portugal
Fernando Cruz Gonçalves, ENIDH, Portugal
Maria Manuela Martins Batista, ENIDH, Portugal

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 25
The Contribution of Data Science Applied to Customer Relationship Management: A Systematic Literature Review 522
Dora Maria Simões, University of Aveiro, Portugal

Chapter 26
Data Science in Economics and Business: Roots and Applications 544
Mara Madaleno, GOVCOPP, DEGEIT, Portugal & University of Aveiro, Portugal
João Lourenço Marques, GOVCOPP, DCSPT, Portugal & University of Aveiro, Portugal
Muhammad Tufail, National University of Sciences and Technology (NUST), Pakistan

Chapter 27
A Historical Review of Immersive Storytelling Technologies: New Uses of AI, Data Science, qnd User Experience in Virtual Worlds 569
Hector Puente Bienvenido, Universidad Complutense de Madrid, Spain
Borja Barinaga, Universidad Francisco de Vitoria Madrid, Spain
Jorge Mora-Fernandez, University of California, San Diego, USA

Chapter 28
Artificial Intelligences Are Subsets of Human Beings, Mainly in Fictions and Films: Challenges and Opportunities 598
Nandini Sen, Heriot Watt University, UK

Chapter 29
Towards Ethical Neuromarketing 2.0 Based on Artificial Intelligence 619
Elodie Attié, Capgemini Engineering, France
Solène Le Bars, Altran Lab, France
Ilhem Quenel, Capgemini Engineering, France

Chapter 30
Responsible Machine Learning for Ethical Artificial Intelligence in Business and Industry 639
Deepak Saxena, Birla Institute of Technology and Science, Pilani, India
Markus Lamest, Trinity College Dublin, Ireland
Veena Bansal, Indian Institute of Technology, Kanpur, India

Compilation of References xxxv

About the Contributors cxix

Index cxxxvi

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Detailed Table of Contents

Preface xxvii

Acknowledgment xxxiv

Section 1
Strategy

Chapter 1
Port Dada Integration: Opportunities for Optimization and Value Creation 1
José Luís Cacho, Port of Sines, Portugal
Adalberto Tokarski, National Waterway Transportation Agency, Brazil
Elizabete Thomas, National Waterway Transportation Agency, Brazil
Valentina Chkoniya, University of Aveiro, Portugal

The maritime supply chain is growing in complexity. Ports are at the crossroads of many activities, modes, and stakeholders, and are actively becoming digital hubs. Today, digital and physical connectivity go hand in hand. The port could benefit from taping the opportunities arising from digitalization and data integration since it helps to leverage external knowledge, engage stakeholders, create new decision-making anchors, lower the risk of certain investments, boost productivity and cut costs, and accelerate greening and digital transition, generating possibilities for just-in-time operations and optimizations. The chapter aims to apprehend the use of data science in the port sector. The state of the art in Brazil and Portugal are different. Even inside Brazil, there is no homogeneity of ports in the usage of digital infrastructure, cloud computing, or artificial intelligence. The existing inequalities hinder general cooperation between nations but, at the same time, reveal opportunities to approach specific nodes in the international supply chain.

Chapter 2
A Software Engineering Perspective on Building Production-Ready Machine Learning Systems..... 23
Petra Heck, Fontys University of Applied Sciences, The Netherlands
Gerard Schouten, Fontys University of Applied Sciences, The Netherlands
Luís Cruz, Delft University of Technology, The Netherlands

This chapter discusses how to build production-ready machine learning systems. There are several challenges involved in accomplishing this, each with its specific solutions regarding practices and tool support. The chapter presents those solutions and introduces MLOps (machine learning operations, also called machine learning engineering) as an overarching and integrated approach in which data engineers, data scientists, software engineers, and operations engineers integrate their activities to implement

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

validated machine learning applications managed from initial idea to daily operation in a production environment. This approach combines agile software engineering processes with the machine learning-specific workflow. Following the principles of MLOps is paramount in building high-quality production-ready machine learning systems. The current state of MLOps is discussed in terms of best practices and tool support. The chapter ends by describing future developments that are bound to improve and extend the tool support for implementing an MLOps approach.

Chapter 3
Multivariate Sustainability Profile of Global Fortune 500 Companies Using GRI-G4 Database 55
Mónica Jiménez-Hernández, Universidad de Colima, Mexico
Purificación Vicente-Galindo, University of Salamanca, Spain
Nathalia Tejedor-Flores, Universidad Tecnológica de Panamá, Panama
Adelaide Freitas, University of Aveiro, Portugal
Purificación Galindo, University of Salamanca, Spain

The main objective of this research is to find the sustainability gradients of Global Fortune 500 companies and sort them as a function of economic, environmental, and social components using multivariate statistical methods to establish the foundations for better knowledge of the trends and sustainability reporting habits. A combined approach, comprising principal coordinates analysis (PCoA) and logistic regression model (LRM), is proposed to build an external logistics biplot (ELB). Moreover, HJ-Biplot and parallel coordinates are applied. This chapter helps to understand why many companies view their corporate social responsibility (CSR) reports as a way to guarantee the credibility of the published information. In particular, based on the Global Reporting Initiative, the sustainability gradients of the Global Fortune 500 companies are obtained and statistically exploited to analyze how the companies can make improvements in terms of sustainability.

Chapter 4
Examining the Evolution of E-Government Development of Nations Through Machine Learning Techniques 85
Niguissie Mengesha, Brock University, Canada
Anteneh Ayanso, Brock University, Canada

Several initiatives have tried to measure the efforts nations have made towards developing e-government. The UN E-Government Development Index (EGDI) is the only global report that ranks and classifies the UN Member States into four categories based on a weighted average of normalized scores on online service, telecom infrastructure, and human capital. The authors argue that the EGDI fails in showing the efforts of nations over time and in informing nations and policymakers as to what and from whom to draw policy lessons. Using the UN EGDI data from 2008 to 2020, they profile the UN Member States and show the relevance of machine learning techniques in addressing these issues. They examine the resulting cluster profiles in terms of theoretical perspectives in the literature and derive policy insights from the different groupings of nations and their evolution over time. Finally, they discuss the policy implications of the proposed methodology and the insights obtained.

Chapter 5
Artificial Intelligence as Driver for SME Competitiveness 108
Nicola Del Sarto, Scuola Superiore Sant'Anna, Italy
Andrea Piccaluga, Scuola Superiore Sant'Anna, Italy

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Artificial intelligence is profoundly changing the way in which companies compete and do business. In particular, artificial intelligence can represent a very interesting opportunity for small and medium-sized enterprises, which are constantly looking for new technologies to be able to remain competitive in a turbulent market. However, research exploring how SMEs may successfully adopt artificial intelligence technology are missing. To address this gap, the authors reviewed the literature on artificial intelligence and identified four key features that SMEs need to consider when implementing this technology represented by people, processes, products, and customers. After that they described four cases of Italian SMEs which have adopted successfully this new technology taking into account one of the four dimensions. The study contributes to the literature on artificial intelligence and SMEs and may be helpful for managers who want to adopt this technology within their company.

Chapter 6

Big Data Is Decision Science: The Case of COVID-19 Vaccination 126
Jacques R. J. Bughin, Université libre de Bruxelles, Belgium
Michele Cincera, Université libre de Bruxelles, Belgium
Dorota Reykowska, NEUROHM, Poland
Rafał Ohme, WSB University in Torun, Poland

Data science has been proven to be an important asset to support better decision making in a variety of settings, whether it is for a scientist to better predict climate change for a company to better predict sales or for a government to anticipate voting preferences. In this research, the authors leverage random forest (RF) as one of the most effective machine learning techniques using big data to predict vaccine intent in five European countries. The findings support the idea that outside of vaccine features, building adequate perception of the risk of contamination, and securing institutional and peer trust are key nudges to convert skeptics to get vaccinated against COVID-19. What machine learning techniques further add beyond traditional regression techniques is some extra granularity in factors affecting vaccine preferences (twice more factors than logistic regression). Other factors that emerge as predictors of vaccine intent are compliance appetite with non-pharmaceutical protective measures as well as perception of the crisis duration.

Chapter 7

The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience..... 151
Elissa Moses, Columbia University, USA & BrainGroup Global, USA
Kimberly Rose Clark, Dartmouth College, USA & Bellwether Citizen Response, USA
Norman J. Jacknis, Columbia University, USA & Intelligent Community Forum, USA

This chapter summarizes the role that artificial intelligence and machine learning (AI/ML) are expected to play at every stage of advertising development, assessment, and execution. Together with advances in neuroscience for measuring attention, cognitive processing, emotional response, and memory, AI/ML have advanced to a point where analytics can be used to identify variables that drive more effective advertising and predict enhanced performance. In addition, the cost of computation has declined, making platforms to apply these tools much less expensive and within reach. The authors then offer recommendations for 1) understanding the clients/customers and users of the products and services that will be advertised, 2) aiding creativity in the process of designing advertisements, 3) testing the impact of advertisements, and 4) identifying the optimum placement of advertisements.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 2
Action

Chapter 8
Retail Customer Native Baskets Creation: A Must-Do Into Measuring Price Indices 168
Liliana Bernardino Matias, Sonae MC, Portugal
Ana Costa Freitas, Sonae MC, Portugal
Filipe Fernandes Miranda, Sonae MC, Portugal

Standard approaches to measure price index on a specific retailer against competitors tend to be overly optimistic. The authors propose a methodology that deflects from analyzing solely the price of all products effectively bought by customers by taking into consideration their prior shopping intention. Loyalty schemes are game changers when measuring price index. The study proposes a two-step methodology that comprises on building the customer native baskets and allocating the prices collected to each customer. Under the assumption that a price sensitive customer will not buy a poorly priced product compared to the competitors in that specific period, the authors present an innovative methodology resulting in a more realistic price index metric.

Chapter 9
Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence 192
M Dolores Méndez-Aparicio, Open University of Catalonia (UOC), Spain
María Pilar Martínez-Ruiz, University of Albacete, Spain
Alicia Izquierdo-Yusta, University of Burgos, Spain
Ana Isabel Jiménez-Zarco, Open University of Catalonia (UOC), Spain

Data science has established itself as a discipline of study in the new technological paradigm of the 21st century. Customer data and metadata are not only unique opportunities for companies that are interested in collecting, processing, and elaborating useful information for customer-centric business management. They are also starting to feed other disciplines, such as industry and computer engineering, which are developing new ways of organising information and intelligent response, where significant technological advances in facial, sensory, text, voice, and image recognition are constantly emerging, with response capabilities ever closer to human thinking. From this continuous interaction emerges a new asset for companies: knowledge. Thus, science applied to data is consolidating as an opportunity for companies in customer loyalty and retention by being able to include a relevant aspect in human relations such as warmth and empathy, basis of marketing 5.0, oriented once again towards human relations, from a virtual existence and total digitalisation.

Chapter 10
Using Intelligent Text Analysis of Online Reviews to Determine the Main Factors of Restaurant Value Propositions 223
Elizaveta Fainshtein, National Research University Higher School of Economics, Russia
Elena Serova, National Research University Higher School of Economics, Russia

This chapter discusses the sentiment classification of text messages containing customer reviews of an online restaurant service system using machine-learning methods, in particular text mining and multivariate text sentiment analysis. The study determines the structure of value proposition factors based

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

on online restaurant reviews on TripAdvisor, collecting information on consumer preferences and the restaurant services in St. Petersburg (Russia) quality assessment and examines the influence of service format and reviews tonality on ratings restaurants factors. The service format context is proposed as the main attribute influencing the formation of the restaurant business value proposition and of relevance for online reviews. The results showed the key factors in the study of the sentiment were cuisine and dishes, reviews and ratings, and targeted search. MANOVA analysis represented that for special offers and features, reviews and ratings, factors and quantitative star ratings influenced the negative and positive sentiment of online reviews significantly.

Chapter 11
Time Series Forecasting in Retail Sales Using LSTM and Prophet 241
Clony Junior, IEETA, University of Aveiro, Portugal
Pedro Gusmão, IEETA, University of Aveiro, Portugal
José Moreira, DETI, University of Aveiro, Portugal
Ana Maria M. Tome, DETI, University of Aveiro, Portugal

Data science highlights fields of study and research such as time series, which, although widely explored in the past, gain new perspectives in the context of this discipline. This chapter presents two approaches to time series forecasting, long short-term memory (LSTM), a special kind of recurrent neural network (RNN), and Prophet, an open-source library developed by Facebook for time series forecasting. With a focus on developing forecasting processes by data mining or machine learning experts, LSTM uses gating mechanisms to deal with long-term dependencies, reducing the short-term memory effect inherent to the traditional RNN. On the other hand, Prophet encapsulates statistical and computational complexity to allow broad use of time series forecasting, prioritizing the expert's business knowledge through exploration and experimentation. Both approaches were applied to a retail time series. This case study comprises daily and half-hourly forecasts, and the performance of both methods was measured using the standard metrics.

Chapter 12
General Model for Metrics Calculation and Behavior Prediction in the Manufacturing Industry:
An Automated Machine Learning Approach 263
Maria João Lopes, Bosch Termotecnologia SA, Portugal & University of Aveiro, Portugal
Eugénio M. Rocha, University of Aveiro, Portugal
Petia Georgieva Georgieva, University of Aveiro, Portugal
Nelson Ferreira, Bosch Termotecnologia SA, Portugal

In this chapter, a comprehensive description of a generic framework aimed at solving various predictive data-driven related use cases, occurring in the manufacturing industry, is provided. The framework is rooted in a general mathematical model so called queue directed graph (QDG). With the aid of QDGs and containerized microservices implementations, the typology of the system is analyzed, and real use cases are explained. The goal is for this framework to be able to be used with all use cases which fit in this typology. As an example, a data generation distribution model is proposed, the parameter stability and predictive robustness are studied, and automated machine learning approaches are discussed to predict the throughput time of products in a manufacturing production line just by knowing the processing time in their first stations.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 13
AI Meaning and Applications in the Consumer Sector of Retailing, Hospitality, and Tourism 291
Sandra Maria Correia Loureiro, ISCTE, BRU, Instituto Universitário de Lisboa, Portugal
Muhammad Ashfaq, Shenzhen University, China
Mariana Oliveira Berga Rodrigues, ISCTE, Instituto Universitário de Lisboa, Portugal

Artificial intelligence (AI) algorithms have been continuously adopted in businesses, such as retailing, hospitality, and tourism. The current chapter aims first to provide a conceptualization of evolution of AI. Second, it gives an overview of AI in business and finally points out examples of applications of AI in retailing, hospitality, and tourism. Third, the chapter offers suggestions for further research. This chapter is devoted to researchers and practitioners. The audience will benefit from an overview of AI meaning and potential research to be developed and also benefit from examples of other business that already employ such AI algorithms.

Chapter 14
Key Factors in the Process of Acceptance and Implementation of Artificial Intelligence in the Hotel Sector 304
Alfonso Infante-Moro, University of Huelva, Spain
Juan C. Infante-Moro, University of Huelva, Spain
Julia Gallardo-Pérez, University of Huelva, Spain

Artificial intelligence is one of the most innovative and trending technologies in the hotel sector. It is transforming the hotel sector into a novel environment in business-client relationships. For this reason, this study seeks to find, through a literary review and a causal study carried out by experts in technologies and hotels, using the methodology of fuzzy cognitive maps, what the determining factors in the decision of hotels to accept and implement artificial intelligence in their hotels and in their services are. The list obtained is made up of the following factors (in order of relevance in this decision): top management support, perceived reliability of the technology, security, relative advantage, support from information systems providers, technological organizational readiness, government pressure or incentives, compatibility, business partner pressure, customer pressure, complexity, pressure from competitors, perceived cost, characteristics of the leader or manager, and size of the company.

Chapter 15
Peruvian Consumer Responses to the Use of Technologies in the Context of COVID-19 323
Yezelia Danira Caceres Cabana, Universidad Nacional de San Agustín de Arequipa, Peru
David Aguilar del Carpio, Universidad Nacional de San Agustín de Arequipa, Peru
Erika Velásquez Chacón, Universidad Católica San Pablo, Peru
Juan Mardonio Rivera Medina, Universidad Continental Perú, Peru

The pandemic caused by COVID-19 has led many states, in an attempt to control the spread of the virus, to decree social immobilization, which meant the introduction of restrictions on the free movement of people and the opening of shops. This led them to seek new marketing channels for purchases. Among these, ICTs have been important. This is the focus of the analysis in this document. Through surveys and interviews, information was obtained, divided into four age groups, which showed that an important part of the population has had to resort to ICTs to acquire goods and/or pay for services. This change in the way of acquiring had different particularities according to the age group analyzed, with a greater change in the oldest group (56 to 74 years old). It can be concluded that the massification of these tools

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

has generated a change in the ways of acquiring products, and this is likely to transcend the pandemic and these channels will be maintained and strengthened in the future.

Chapter 16
Artificial Intelligence in Marketing: A Review of Consumer-AI Interactions..... 342
Özge Sığırcı, Kırklareli University, Turkey

The purpose of this chapter is to shed light on the consumer-AI interaction in the marketplace. By this aim, the chapter uses a literature review approach. The previous literature examining AI from a consumer behavior perspective is reviewed, and the findings are compiled in a meaningful flow. According to the review, we see that the traditional marketplace is shaped by AI from only human-to-human interactions to human-to-AI and AI-to-AI interactions. In this new marketplace, while consumers interact with AI, they gain new experiences and feel positive or negative because of these experiences. Also, they build different relationships with AI, such as servant, master, or partner. Besides these relationships, there are still concerns about AI that are related to privacy, algorithmic biases, consumer vulnerability, unemployment, and ethical decision making.

Chapter 17
When Luxury Vinous-Concept Hotel Meets Premium Wine Brands: An Exploratory Study on Co-Branding..... 366
Lisa Ferraz, University of Aveiro, Portugal
Helena Nobre, University of Aveiro, Portugal
Belém Barbosa, University of Porto, Portugal

Co-branding in the hospitality luxury sector is still understudied in the literature. This study aims at tackling this gap through the analysis of a case of a co-branding strategy between a vinous-concept luxury hotel in Portugal and premium wine brands of domestic producers. Fourteen in-depth interviews with managers of the luxury hotel and wine brand partners supported the exploratory research. This chapter represents a case of qualitative data application to an underestimated topic in the literature from the managers' point of view. The study offers evidence on the benefits for both parties, reasons for adopting co-branding, and partners' selection attributes. The improvement of brand image emerges as one of the main advantages of co-branding with a luxury hotel. Based on the literature review and the interviews with managers, the study proposes a set of hypotheses to be tested in future research. This chapter provides interesting cues for academics and practitioners.

Chapter 18
Data Analysis in the Shipping Industry: eShip Case Study – Problem Statement 381
Marcel Kyas, Reykjavik University, Iceland
Joshua D. Springer, Reykjavik University, Iceland
Jan Tore Pedersen, Marlo as., Norway
Valentina Chkoniya, University of Aveiro, Portugal

This chapter identifies the critical issues that must be addressed to accelerate the digital transition in the chartering market. The maritime industry is one of the pillars of global trade, where change is a constant. Again, shipping is at the cusp of a new era—one driven by data. The authors review the state-of-the-art technology that is useful to automate chartering processes. · The Fourth Industrial Revolution (or Industry 4.0) starts to change the bulk shipping markets leveraging the data flow between industrial

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

processes in the physical and virtual world. · The internet of things accelerates data flow from things in the real world to the virtual world and enables us to control processes in real-time. Machine-to-machine communication, together with artificial intelligence, creates autonomous systems in many areas of production and logistics. Based on the gathered elements, eShip's case study was analyzed, and future steps have been defined for the data analysis in the shipping industry.

Chapter 19

Question Answering Chatbot Using Memory Networks 401
Riddhi Patel, Charotar University of Science and Technology, India
Parth Patel, Charotar University of Science and Technology, India
Chintan M. Bhatt, Charotar University of Science and Technology, India
Mrugendrasinh L. Rahevar, Charotar University of Science and Technology, India

Manipulation of a large amount of data becomes a very tedious task. Hence, the authors took the approach of memory networks for the implementation of the chatbot. Traditionally, the LSTM model was used to implement chatbots and QA systems. But the LSTM failed to store relevant information when given a longer information set. On the contrary, the memory networks have an additional memory component with it. This can help in storing long information for further use which is greatly advantageous for the QA and chatbot systems as compared to LSTM. The authors trained and tested their model over Facebook's bAbi dataset which consists of several tasks and has questions regarding each task to retrieve the accuracy of the model. On the pedestal of that dataset, they have presented the accuracy for every task in their study with memory networks.

Chapter 20

Comparison of Parametric and Non-Parametric Methods to Analyse the Data Gathered by a Likert-Type Scale 414
Tamer Baran, Pamukkale University, Turkey

The aim of this chapter is to reveal whether the results of the analysis of the data obtained using Likert type scales (LTSs) with parametric and non-parametric methods in different response alternative (DRA) numbers will differ in terms of statistical significance. In this respect, the data were obtained from 271 university students with CETSCALE prepared using LTS in five different response alternatives (DRAs). The data were analysed using the one sample t test and Wilcoxon signed rank test. Significant findings of the study in the analysis of the data obtained using midpoint LTSs and with the normal distribution with both parametric and non-parametric methods couldn't be found. Similarly, the data obtained by four response alternative numbers with the normal distribution were analysed by both methods, and the significant findings were revealed. However, the results of the data obtained by six and eight response alternative numbers with parametric methods were found to be statistically significant while their analysis by non-parametric methods did not reveal significant findings.

Chapter 21

Latent Fingerprint Enhancement Based on EDTV Model..... 431
Shadi M. S. Hilles, Istanbul Okan University, Turkey
Abdilahi Deria Liban, Hargeisa University, Somalia
Abdullah M. M. Altrad, Al-Madinah International University, Malaysia
Yousef A. Baker El-Ebiary, Universiti Sultan Zainal Abidin, Malaysia
Mohanad M. Hilles, University College of Applied Science, Palestine

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The chapter presents latent fingerprint enhancement technique for enforcement agencies to identify criminals. There are many challenges in the area of latent fingerprinting due to poor-quality images, which consist of unclear ridge structure and overlapping patterns with structure noise. Image enhancement is important to suppress several different noises for improving accuracy of ridge structure. The chapter presents a combination of edge directional total variation model, EDTV, and quality image enhancement with lost minutia re-construction, RMSE, for evaluation and performance in the proposed algorithm. The result shows the average of three different image categories which are extracted from the SD7 dataset, and the assessments are good, bad, and ugly, respectively. The result of RMSE before and after enhancement shows the performance ratio of the proposed method is better for latent fingerprint images compared to bad and ugly images while there is not much difference with performance of bad and ugly.

Chapter 22
The Business, Societal, and Enterprise Architecture Framework: An Artificial Intelligence-, Data Sciences-, and Big Data-Based Approach 447
Antoine Trad, IBISTM, France

In this chapter, the author presents an artificial intelligence (AI)-based generic concept for decision making using data science. The applied holistic mathematical model for AI (AHMM4AI) focuses on data and access management. A decisive business decision in a business transformation process of a traditional business environment into an automated AI-based business environment is the capacity of the decision-making system and the profile of the business transformation manager (BTM, or simply the manager). The manager and his team are supported by a holistic framework. The role of data science and the needed data modelling techniques are essential for managing data models in a transformation project. For that reason, the development of the big data for AI (BGD4AI) is an essential start.

Chapter 23
The Effect of Population, Public Investments, and Economic Growth on Regional Renewable Energy Consumption in Turkey: Dynamic Panel Data Analysis 491
Özlem Karadağ Albayrak, Kafkas University, Turkey

Making the use of renewable energy sources widespread is of paramount importance for Turkey as for all countries. In this regard, the determiners of renewable energy consumption have been investigated. The effect of determining or factors affecting the use of renewable energy sources on a regional scale to Turkey were examined with different qualitative and quantitative research techniques. In this study, the factors of economic growth, public investments, and population are analyzed by considering regional differences on the consumption of renewable energy resources. The effect of regional economic growth, regional public investments, and regional population on the amount of regional renewable energy consumption were investigated by using panel data of 26 statistical regions of Level-2 classification in the period between 2010-2018 in Turkey. The results obtained by the dynamic panel data analysis concluded that economic growth and public investments at the regional level increased renewable energy consumption while the population growth decreased.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 3
Reflection

Chapter 24
Education for the Digital Industry: Opportunities and Challenges of Experience-Based Expertise and Open Innovation 506
Valentina Chkoniya, University of Aveiro, Portugal
Fernando Cruz Gonçalves, ENIDH, Portugal
Maria Manuela Martins Batista, ENIDH, Portugal

Education holds the power to transform and enrich the lives of people. In the era of the digital industry, where data is omnipresent in every walk of life and new trends impact society and future jobs, humans continue to evolve through education and developing mechanisms to improve education with data science in the heart of it. This chapter demonstrates that experience-based expertise and open innovation must be understood as a single process, where living labs that involve academies and enterprises create unique conditions for society's progress. There is a trinomial relationship between academy, society, and industries, which are interestingly far more exploited than the education and research. Effective management of the knowledge and information transferred between open innovation ecosystem partners is crucial. The scientific development of both concepts is an active field in the academic community, and new ideas appear, opening new paths of knowledge transfer methods with knowledge from data.

Chapter 25
The Contribution of Data Science Applied to Customer Relationship Management: A Systematic Literature Review 522
Dora Maria Simões, University of Aveiro, Portugal

In the face the contemporary world lives, and the consequent data produced at an unprecedented speed through digital media platforms, the data are nowadays called the new global currency. It raises numerous opportunities to improve outcomes in businesses, namely at the level of customer relationship management (CRM) strategies and their systems. Nevertheless, how analytics can be applied and support the customer relationship processes seems unclear for academics and industries. To better connect customer relationship processes needs and what data science analytics can offer, this chapter presents a systematic literature review around the concepts, tools, and techniques behind this field, looking particularly on customer acquisition and customer retention in businesses. The outcomes highlight that academic researcher works in this field are very scare and recent. Searching the Scopus and Web of Science databases resulted in only 12 documents from 2013 to 2020, eight of them published in the last two years.

Chapter 26
Data Science in Economics and Business: Roots and Applications 544
Mara Madaleno, GOVCOPP, DEGEIT, Portugal & University of Aveiro, Portugal
João Lourenço Marques, GOVCOPP, DCSPT, Portugal & University of Aveiro, Portugal
Muhammad Tufail, National University of Sciences and Technology (NUST), Pakistan

Economics and business are a great background for data science provided econometricians and data scientists are sets with an intersection, although remaining unknown. In econometrics, data mining is somewhat a monstrous word, a field that traditionally seeks causal inference and results in interpretability. When we go deeper into what data science usually is, the boundaries between more traditional econometrics and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

even statistics and the hip and cool machine learning become shorter. In economics and business, we find examples and applications of simple and advanced data science techniques. This chapter intends to provide state-of-the-art data science applications in economics and business. The review and bibliometric analysis are limited to the research articles published through Elsevier Scopus. Results allowed the authors to conclude that despite the number of already existent research, a lot more remains to be explored joining both fields of knowledge, data since, and economics and business. This analysis allowed the authors to identify further possible avenues of research critically.

Chapter 27
A Historical Review of Immersive Storytelling Technologies: New Uses of AI, Data Science, qnd User Experience in Virtual Worlds 569
Hector Puente Bienvenido, Universidad Complutense de Madrid, Spain
Borja Barinaga, Universidad Francisco de Vitoria Madrid, Spain
Jorge Mora-Fernandez, University of California, San Diego, USA

This chapter is focused on describing the history and the current relevance of user experience (UX) techniques that combine data science and AI in the research field of interactive and immersive storytelling, including virtual and augmented realities. It initially presents a brief history of interactive storytelling, video games, VR and AR, AI and data science, and the user experience (UX) techniques used in those areas. Later, the chapter describes the UX techniques in depth, using AI and data science that work best and are more useful for testing interactive media products, describing examples of its applications briefly. Finally, the chapter presents conclusions in relationship with utopias and dystopias regarding the future use of UX, AI, and data science in several areas such as edutainment, social media, media arts, and business, among others.

Chapter 28
Artificial Intelligences Are Subsets of Human Beings, Mainly in Fictions and Films: Challenges and Opportunities 598
Nandini Sen, Heriot Watt University, UK

This chapter aims to create new knowledge regarding artificial intelligence (AI) ethics and relevant subjects while reviewing ethical relationship between human beings and AI/robotics and linking between the moral fabric or the ethical issues of AI as used in fictions and films. It carefully analyses how a human being will love robot and vice versa. Here, fictions and films are not just about technology but about their feelings and the nature of bonding between AIs and the human race. Ordinary human beings distrust and then start to like AIs. However, if the AI becomes a rogue as seen in many fictions and films, then the AI is taken down to avoid the destruction of the human beings. Scientists like Turing are champions of robot/AI's feelings. Fictional and movie AIs are developed to keenly watch and comprehend humans. These actions are so close to empathy they amount to consciousness and emotional quotient.

Chapter 29
Towards Ethical Neuromarketing 2.0 Based on Artificial Intelligence 619
Elodie Attié, Capgemini Engineering, France
Solène Le Bars, Altran Lab, France
Ilhem Quenel, Capgemini Engineering, France

Eighty percent of consumer behaviors and purchases rely on subconscious processes. The use of

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

neuromarketing tools to study consumer behavior is not clear, notably regarding its practices and intentions toward consumers. This chapter aims to understand how neuromarketing can explain consumer behavior thanks to Neuromarketing 2.0 tools, how companies can manage the collected data in a responsible way and build a neuroethical charter to regulate the way companies use it. Most companies choose to not communicate about it when they use neuromarketing tools, and therefore, this chapter aims to pave the way towards solutions and recommendations and democratize its use by making Neuromarketing 2.0 more responsible and ethical.

Chapter 30
Responsible Machine Learning for Ethical Artificial Intelligence in Business and Industry 639
Deepak Saxena, Birla Institute of Technology and Science, Pilani, India
Markus Lamest, Trinity College Dublin, Ireland
Veena Bansal, Indian Institute of Technology, Kanpur, India

Artificial intelligence (AI) systems have become a new reality of modern life. They have become ubiquitous to virtually all socio-economic activities in business and industry. With the extent of AI's influence on our lives, it is an imperative to focus our attention on the ethics of AI. While humans develop their moral and ethical framework via self-awareness and reflection, the current generation of AI lacks these abilities. Drawing from the concept of human-AI hybrid, this chapter offers managerial and developers' action towards responsible machine learning for ethical artificial intelligence. The actions consist of privacy by design, development of explainable AI, identification and removal of inherent biases, and most importantly, using AI as a moral enabler. Application of these action would not only help towards ethical AI; it would also help in supporting moral development of human-AI hybrid.

Compilation of References xxxv

About the Contributors cxix

Index cxxxvi

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Preface

"Doing the same thing repeatedly, and expecting different results is the definition of insanity."

– Albert Einstein (1879 - 1955)

Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry brings a new unique and global perspective on the subject of 100 highly professional and experienced researchers and practitioners from 27 different countries: Belgium, Brazil, Canada, China, Colombia, France, Georgia, Iceland, India, Ireland, Italy, Malaysia, Mexico, Nederland, Norway, Pakistan, Palestine, Panama, Peru, Poland, Portugal, Russia, Somaliland, Spain, Turkey, United Kingdom, and the USA.

The contemporary world lives on the data produced at an unprecedented speed through social networks and the internet of things (IoT). Data has been called the new global currency, and its rise is transforming entire industries, providing a wealth of opportunities. Applied data science research is necessary to derive useful information from big data for effective and efficient utilization to solve real-world problems. A broad analytical set allied with solid business logic is fundamental in today's corporations. Organizations work to obtain a competitive advantage by analyzing the data produced within and outside their corporate limits to support their decision-making processes. This book aims to provide an overview of the concepts, tools, and techniques behind the fields of data science and artificial intelligence (AI) applied to business and industries.

Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry offers a compilation of 30 chapters presenting a holistic approach and is organized in three sections: "Strategy," "Action," and "Reflection." The spread of this orientation has meant preceding perspective, reminding that Actions have a greater chance of success if Strategies guide them. Furthermore, Reflection on the Action is vital for adjustments in the Strategy.

STRATEGY

The seven chapters that compose the first section of the book introduce the Strategic Perspective on Applied Data Science and Artificial Intelligence in Business and Industry.

In the first chapter, "Port Dada Integration: Opportunities for Optimization and Value Creation," José Luís Cacho, Adalberto Tokarski, Elizabete Thomas, and Valentina Chkoniya bring the academy, Port of Sines (Portugal), and National Waterway Transportation Agency from Brazil together. The chapter aims

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to apprehend the use of data science in the port sector. It is imperative in the Maritime Supply Chain that is growing in complexity. Ports are at the crossroads of many activities, modes, and stakeholders, actively becoming a digital hub. Today digital and physical connectivity go hand in hand. The Port could benefit from taping the opportunities arising from digitalization and data integration since it helps to leverage external knowledge, engage stakeholders, create new decision-making anchors, lower the risk of certain investments, boost productivity and cut costs, accelerate greening and digital transition, generating possibilities for just-in-time operations and optimizations.

The second chapter, "A Software Engineering Perspective on Building Production-Ready Machine Learning Systems," by Petra Heck, Gerard Schouten, and Luis Cruz, discusses building production-ready machine learning systems. There are several challenges involved in accomplishing this, each with its specific solutions regarding practices and tool support. The chapter presents those solutions and introduces MLOps as an overarching and integrated approach. Data engineers, data scientists, software engineers, and operations engineers integrate their activities to implement validated machine learning applications managed from initial ideas to daily operation in a production environment.

In the third chapter, "Multivariate Sustainability Profile of Global Fortune 500 Companies Using GRI-G4 Database," international mathematicians team composed by Mónica Jiménez-Hernández, Purificación Vicente-Galindo, Nathalia Diazibeth Tejedor-Flores, Adelaide Freitas, and Purificación Galindo help to understand why many companies view their Corporate Social Responsibility reports as a way to guarantee the credibility of the published information. Moreover, explain how to find the sustainability gradients and sort them as a function of economic, environmental and, social components, using multivariate statistical methods, to establish the foundations for better knowledge of the trends and sustainability reporting habits.

The fourth chapter, "Examining the Evolution of E-Government Development of Nations Through Machine Learning Techniques," by Niguissie Mengesha and Anteneh Ayanso, argues that the EGDI fails to show nations' efforts over time and in informing nations and policymakers as to what and from whom to draw policy lessons. The authors show the relevance of machine learning techniques in addressing these issues and discuss the policy implications of their proposed methodology and the insights obtained.

The fifth chapter, "Artificial Intelligence as Driver for SMEs Competitiveness," by Nicola Del Sarto and Andrea Piccaluga, shows the importance of exploring how SMEs may successfully adopt Artificial intelligence technology and describe four cases of Italian SMEs. The authors base their analyses on four dimensions represented by People, Processes, Products, and Customers. The study intends to help managers to adopt this technology within their company.

In the sixth chapter, "Big Data Is Decision Science: The Case of COVID-19 Vaccination," Jacques RJ Bughin, Michele Cincera, Dorota Reykowska, and Rafał Ohme suggest that Data science has been proven to be an essential asset to support better decision-making in a variety of settings. Whether it is for a scientist to predict climate change better, a company to better predict sales, or a government to anticipate voting preferences. In this research, authors leverage Random Forest as one of the most effective machine learning techniques using big data to predict vaccine intent in five European countries.

Following the neuromarketing perspective, the seventh chapter, "The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience," was written by Elissa Moses, Kimberly Rose Clark, and Norman J. Jacknis. The authors summarize the role that AI/ML are expected to play at every stage of advertising development, assessment, and execution. With advances in neuroscience for measuring attention, cognitive processing, emotional response, and memory, AI/ML have advanced to a point where analytics can be used to identify variables that drive

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

more effective advertising and predict enhanced performance. In addition, the cost of computation has declined, making platforms to apply these tools much less expensive and within reach. The authors also offer recommendations for a) understanding the clients/customers and users of the products and services that will be advertised, b) aiding creativity in designing advertisements, c) testing the impact of advertisements, and d) identifying the optimum placement of advertisements.

ACTION

Section 2 of this book (Action) addresses that Actions have a greater chance of success if Strategies guide them.

The following 16 chapters help to understand how to transform Strategy into Action. They show that Applied Data science is focused on Making Sense of Data. Furthermore, AI is a collection of mathematical algorithms that make computers understand relationships between different types and pieces of data. This knowledge of connections could be utilized to come to conclusions or make decisions. The hardest part of running a successful data-driven business intelligence system has the Strategies to Act on the insights.

The eighth chapter, "Retail Customers' Native Baskets Creation: A Must-Do Into Measuring Price Indexes," by Liliana Bernardino Matias, Ana Costa Freitas, and Filipe Fernandes Miranda, gives a practical perspective of the most prominent Portuguese retailer, Sonae MC, and proposes a methodology that deflects from analyzing solely the price of all products effectively bought by customers by taking into consideration their prior shopping intention. Loyalty schemes are game-changers when measuring price index, and standard approaches to measure price index on a specific retailer against competitors tend to be overly optimistic. The study proposes a two-step methodology comprising building the customer native baskets and allocating the prices collected to each customer. Under the assumption that a price-sensitive customer will not buy a poorly priced product compared to the competitors in that specific period, the authors present an innovative methodology resulting in a more realistic price index metric.

In the ninth chapter, "Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence," Dolores Méndez-Aparicio, María Pilar Martínez-Ruiz, Alicia Izquierdo-Yusta, and Ana Isabel Jiménez-Zarco focus on customer data and metadata. There is a unique opportunity for companies interested in collecting, processing, and elaborating useful information for customer-centric business management. Other disciplines are developing new ways of organizing information and intelligent response. Significant technological advances in facial, sensory, text, voice, and image recognition are constantly emerging, with response capabilities closer to human thinking. From this continuous interaction emerges a new asset for companies: knowledge.

In the tenth chapter, "Using Intelligent Text Analysis of Online Reviews to Determine the Main Factors of Restaurant Value Propositions," Elizaveta Fainshtein and Elena Serova discuss the sentiment classification of text messages containing customer reviews of an online restaurant service system, using machine-learning methods, in particular, text mining and multivariate text sentiment analysis. The study determines the value proposition factors based on online restaurant reviews on TripAdvisor, collecting information on consumer preferences and restaurant services in St. Petersburg (Russia) quality assessment and examining the influence of service format and reviews tonality on rating restaurants factors. The results showed that the critical factors in the sentiment study were cuisine and dishes, reviews and ratings, and targeted search.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the eleventh chapter, "Time Series Forecasting in Retail Sales Using LSTM and Prophet," Clony Junior, Pedro Gusmão, José Moreira, and Ana Maria M. Tome present two approaches to time series forecasting, Long Short-Term Memory, a special kind of Recurrent Neural Network, and Prophet, an open-source library developed by Facebook for time series forecasting. With a focus on developing forecasting processes by data mining or machine learning experts, LSTM uses gating mechanisms to deal with long-term dependencies, reducing the short-term memory effect inherent to the traditional RNN. On the other hand, Prophet encapsulates statistical and computational complexity to allow broad use of time series forecasting, prioritizing the expert's business knowledge through exploration and experimentation. Both approaches were applied to a retail time series. This case study comprises daily and half-hourly forecasts and the performance of both methods was measured using the standard metrics.

The twelfth chapter, "General Model for Metrics Calculation and Behavior Prediction in the Manufacturing Industry: An Automated Machine Learning Approach," by Maria João Lopes, Eugénio M. Rocha, Petia Georgieva Georgieva, Nelson Ferreira, shows an example of the applied data science in Bosch Termotecnologia SA. Authors provide a comprehensive description of a generic framework aimed at solving various predictive data-driven related use cases, occurring in the manufacturing industry. As an example, a data generation distribution model is proposed, parameter stability and predictive robustness are studied, and automated machine learning approaches are discussed to predict the throughput time of products in a manufacturing production line just by knowing the processing time in their first stations.

In the thirteenth chapter, "AI Meaning and Applications in the Consumer Sector of Retailing, Hospitality, and Tourism," Sandra Maria Correia Loureiro, Muhammad Ashfaq, and Mariana Oliveira Berga Rodrigues provide a conceptualization of the evolution of AI. The authors' overview AI in business and point out examples of AI applications in retailing, hospitality, and tourism.

Following the topic, in the fourteenth chapter, "Key Factors in the Process of Acceptance and Implementation of Artificial Intelligence in the Hotel Sector," Alfonso Infante-Moro, Juan C. Infante-Moro, and Julia Gallardo-Pérez show that AI is one of the most innovative and trending technologies in the hotel sector, which is transforming the hotel sector into a novel environment in business-client relationships. For this reason, this study seeks to find, through a literary review and a causal study carried out by experts in technologies and hotels, using the methodology of fuzzy cognitive maps, what are the determining factors in the decision of hotels to accept and implement the Artificial Intelligence in their hotels and their services.

The fifteenth chapter, "Peruvian Consumer Responses to The Use of Technologies in the Context of COVID-19," by Yezelia Danira Caceres Cabana, David Aguilar del Carpio, Erika Velásquez Chacón, and Juan Mardonio Rivera Medina, gives a transformational perspective of marketing channels for purchases. ICTs are the focus of the analysis. Authors conclude that the massification of these tools has generated a change in the ways of acquiring products, which is likely to transcend the pandemic, and these channels will be maintained and strengthened in the future.

In the sixteenth chapter, "Artificial Intelligence in Marketing: A Review of Consumer-AI Interactions," Özge Sığırcı analyzes the consumer-AI interaction in the marketplace. Besides these relationships, there are still concerns about AI-related privacy, algorithmic biases, consumer vulnerability, unemployment, and ethical decision-making.

In the seventeenth chapter, "When Luxury Vinous-Concept Hotel Meets Premium Wine Brands: An Exploratory Study on Co-Branding," Lisa Ferraz, Helena Nobre, and Belém Barbosa analyze a case of a co-branding strategy between a vinous-concept luxury hotel in Portugal and premium wine brands

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of domestic producers. The study offers evidence on the benefits for both parties, reasons for adopting co-branding, and partners' selection attributes.

The eighteenth chapter is devoted to "Data Analysis in the Shipping Industry: eShip Case Study – Problem Statement" and written by Marcel Kyas, Joshua D Springer, Jan Tore Pedersen, and Valentina Chkoniya. This chapter identifies the critical issues that must be addressed to accelerate the digital transition in the chartering market. The maritime industry is one of the pillars of global trade, where change is a constant. Again, shipping is at the cusp of a new era—one driven by data. eShip's case study was analyzed, and future steps have been defined for the data analysis in the shipping industry.

In the nineteenth chapter, "Question Answering Chatbot Using Memory Networks," Riddhi Patel, Parth Patel, Chintan M Bhatt, and Mrugendrasinh L Rahevar trained and tested their model over Facebook's bAbi dataset, which consists of several tasks and has questions regarding each task to retrieve the accuracy of the model. On the pedestal of that dataset, they have presented the accuracy for every task in their study with memory networks.

The twentieth chapter, "Comparison of Parametric and Non-Parametric Methods to Analyse the Data Gathered by Likert Type Scale," by Tamer Baran, examines whether the results of the analysis of the data obtained using LTSs with parametric and non-parametric methods in different response alternative numbers will differ in terms of statistical significance.

In the twenty-first chapter, "Latent Fingerprint Enhancement Based on EDTV Model," Shadi M S Hilles, Abdilahi Deria Liban, Abdullah M. M Altrad, Yousef A. Baker El-Ebiary, and Mohanad M Hilles underline challenges in the area of latent fingerprint due to poor-quality image which consist unclear ridge structure and overlapping patterns with structure noise, where is a latent fingerprint in image enhancement is essential to suppress several different noises for improved accuracy of ridge structure, presents the combination of edge directional total variation model EDTV and quality image enhancement with lost minutia re-construction, RMSE for evaluation and performance in the proposed algorithm.

The twenty-second chapter, "The Business, Societal and Enterprise Architecture Framework: An Artificial Intelligence, Data Sciences, and Big Data-Based Approach," by Antoine Trad, presents an AI-based generic concept for decision making using data science. The Applied Holistic Mathematical Model for AI (AHMM4AI) focuses on data and access management, and the author demonstrates that the development of the Big Data for AI (BGD4AI) is an essential jumpstart.

In the twenty-third chapter, "The Effect of Population, Public Investments, and Economic Growth on Regional Renewable Energy Consumption in Turkey: Dynamic Panel Data Analysis," Özlem Karadağ Albayrak investigates determiners of renewable energy consumption. In this study, the factors of economic growth, public investments, and population are analyzed by considering regional differences in the consumption of renewable energy resources. The effect of regional economic growth, regional public investments, and regional population on regional renewable energy consumption was investigated.

REFLECTION

Section 3 (Reflection) addresses several dimensions uncovered in this book. Actionable insight is an effect for which the root cause can be directly and inarguably derived with enough specificity to know how to Action a Change. Knowing how to identify an actionable insight is only the first step, in any case. Data science is about decision-making in the increasingly data-driven world we live in today. It has employed great efforts in developing advanced analytics, improving data models, and cultivating new

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

algorithms. However, there are organizational and socio-technical challenges that arise when executing a Data Science project.

The last seven chapters point to three layers to consider in this field: knowledge, practice, and ethics.

In the twenty-fourth chapter, "Education for The Digital Industry: Opportunities and Challenges of Experience-Based Expertise and Open Innovation," Valentina Chkoniya, Fernando Cruz Gonçalves, and Maria Manuela Martins Batista suggest in the era of the Digital Industry, where data is omnipresent in every walk of life, and new trends impact society and future jobs, humans continue to evolve through education and developing mechanisms to improve education with data science in the heart of it. Effective management of the knowledge and information transferred between open innovation ecosystem partners is crucial. The scientific development of both concepts is an active field in the academic community, and new ideas appear, opening new paths of knowledge transfer methods with knowledge from data.

In the twenty-fifth chapter, "The Contribute of Data Science Applied to Customer Relationship Management: A Systematic Literature Review," Dora Maria Simões defends that in the face the contemporary world lives, and the consequent data produced at an unprecedented speed through digital media platforms, the data are nowadays called the new global currency. It raises numerous opportunities to improve outcomes in businesses, namely a level of customer relationship management (CRM) strategies and its systems.

In the twenty-sixth chapter, "Data Science in Economics and Business: Roots and Applications," Mara Madaleno, João Lourenço Marques, and Muhammad Tufail show that economics and business are excellent backgrounds for data science provided econometricians and data scientists are sets with an intersection, although remaining unknown. Reflecting on what data science usually is.

The twenty-seventh chapter, "A Historical Review of Immersive Storytelling Technologies: New Uses of AI, Data Science, and User Experience in Virtual Worlds," by Hector Puente Bienvenido, Borja Barinaga, and Jorge Mora-Fernandez focus on describing the history and the current relevance of UX techniques that combine data science and AI in the research field of interactive and immersive storytelling, including Virtual and Augmented Realities. It initially presents a brief history of interactive storytelling, video games, VR and AR, AI and data science, and the UX techniques used in those areas. Later, the chapter describes the UX techniques in-depth, using AI and data science that work best and are more helpful in testing interactive media products, describing examples of its applications briefly. Finally, the chapter presents conclusions in relationship with utopias and dystopias regarding the future use of UX, AI, and data science in several areas such as edutainment, social media, media arts, and business, among others.

In the twenty-eighth chapter, "Artificial Intelligences Are Subsets of Human Beings, Mainly in Fictions and Films: Challenges and Opportunities," anthropologist Nandini Sentries tries to create new knowledge regarding AI ethics. The author also revives the ethical relationship between human beings and AI/robotics and linking between the moral fabric or the ethical issues of AI.

The twenty-ninth chapter, "Towards Ethical Neuromarketing 2.0 Based on Artificial Intelligence," by Elodie Attié, Solène Le Bars, and Ilhem Quesnel, explains the perspective of Capgemini Engineering and Altran Lab on the use of neuromarketing tools to study consumer behavior. This chapter aims to understand how neuromarketing can explain consumer behavior thanks to neuromarketing 2.0 tools, how companies can responsibly manage the collected data, and build a neuroethical charter to regulate the way companies use it.

Finally, in the thirtieth chapter, "Responsible Machine Learning for Ethical Artificial Intelligence in Business and Industry," Deepak Saxena, Markus Lamest, and Veena Bansal examine AI systems as

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

a new reality of modern life. They have become ubiquitous to virtually all socio-economic activities in business and industry, while humans develop their moral and ethical framework via self-awareness and reflection. Based on the concept of human-AI hybrid, this chapter offers managerial and developers' action towards responsible machine learning for ethical artificial intelligence.

As the book demonstrates, Data Science is all about data. Applied Data science is focused on Making Sense of Data. AI is a collection of mathematical algorithms that make computers understand relationships between different types and pieces of data such that this knowledge of connections could be utilized to come to conclusions or make decisions that could be accurate to a very high degree.

Featuring practice and research on topics such as AI, applied data science, big data, business intelligence, consumer behavior analytics, customer experience, customer response modeling, deep learning, ethics, digital marketing analytics, logistic and supply chain, machine learning, market monitoring, marketing intelligence, and marketing strategy analytics.

Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry is a truly international, interdisciplinary, and integrated vision on the topic of Data Science Applications. They lead to exciting and worthwhile work meant to be used as a basis for further research and as an essential reference source for the practice and those concerned with their contribution to a better world.

Looking at things from a different, global, and holistic perspective, we will see the Most Effective and Innovative Strategies for Sustainable Success!

Valentina Chkoniya
University of Aveiro, Portugal

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Acknowledgment

The editor would like to acknowledge the contribution of all the people involved in this project that accepted embarking on the challenge of bridging the gap between Academy and Enterprise.

The path has been long, and those who prevailed can be proud of the final result of a unique holistic perspective of Research on Applied Data Science and Artificial Intelligence in Business and Industry.

We have made a great team of 100 contributors (authors and editorial advisory board) from 27 countries:

Belgium, Brazil, Canada, China, Colombia, France, Georgia, Iceland, India, Ireland, Italy, Malaysia, Mexico, Nederland, Norway, Pakistan, Palestine, Panama, Peru, Poland, Portugal, Russia, Somaliland, Spain, Turkey, United Kingdom, and the USA.

Thank YOU all for your contribution!

Valentina Chkoniya
University of Aveiro, Portugal

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 1
Strategy

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 1
Port Dada Integration: Opportunities for Optimization and Value Creation

José Luís Cacho
https://orcid.org/0000-0001-7854-871X
Port of Sines, Portugal

Adalberto Tokarski
National Waterway Transportation Agency, Brazil

Elizabete Thomas
National Waterway Transportation Agency, Brazil

Valentina Chkoniya
https://orcid.org/0000-0003-1174-3216
University of Aveiro, Portugal

ABSTRACT

The maritime supply chain is growing in complexity. Ports are at the crossroads of many activities, modes, and stakeholders, and are actively becoming digital hubs. Today, digital and physical connectivity go hand in hand. The port could benefit from taping the opportunities arising from digitalization and data integration since it helps to leverage external knowledge, engage stakeholders, create new decision-making anchors, lower the risk of certain investments, boost productivity and cut costs, and accelerate greening and digital transition, generating possibilities for just-in-time operations and optimizations. The chapter aims to apprehend the use of data science in the port sector. The state of the art in Brazil and Portugal are different. Even inside Brazil, there is no homogeneity of ports in the usage of digital infrastructure, cloud computing, or artificial intelligence. The existing inequalities hinder general co-operation between nations but, at the same time, reveal opportunities to approach specific nodes in the international supply chain.

DOI: 10.4018/978-1-7998-6985-6.ch001

This chapter published as an Open Access Chapter distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The present work is an exploratory study that aims to understand the application of data science in the port sector. The purpose, at first, is to provide information about the state of the art in two selected countries and then is to inform how it is possible to integrate data between ports from different nations to fill existing knowledge gaps. Information is currently generated in distinct regions or points in the supply chain, and it is necessary to identify a way to connect it systematically.

Another goal is to gather information about what would be the gains from these processes. Initially, the automation of routines with the integration of different systems is envisaged, which would result in resource savings, aid to decision-making processes, reduction of waiting time in ports, and the decrease of processes and goods costs. What other port challenges can be faced via data integration? Cooperation with other stakeholders will also be crucial to find workable solutions for issues like greening the shipping sector or the digitalization of the port ecosystem. According to European Sea Ports Organisation (ESPO), by teaming up, ports can leverage external knowledge or lower the risk of particular investments (Deloitte & ESPO, 2021). The study intends to trace a path that, later, can be followed.

After analyzing the trends on supply chains and the uses of machine learning in maritime logistics (which encompass Digital Supply Chain, Digital Ship, and Digital Prots), this chapter describes the current scenario in Brazil and Portugal. Survey the types of information produced by their ports and study the possibility of data exchange between them to identify the potential gains and challenges faced with data science.

This chapter aims to make an important locus in the global supply chain visible: the harbor. Several players operate in it, and it is hard to define, among the myriad of data produced, which ones can be made available and can optimize the system or bring tangible benefits to commercial exchanges. The scope of the present work is, therefore, bold. It is intended, in the end, to name the obstacles, assess their impact on the exchange of goods, and outline strategies to overcome them.

In summary, the authors seek to translate the complex reality of the sector into words and identify possible improvements resulting from data exchange in a systematic way. As a case study, the possibility of integration between selected Portuguese and Brazilian ports will be explored. Subsequently, future challenges in this context and possible lines of research will be outlined.

Background

There are more than 7500 ports worldwide that together accommodate the vast majority of world trade volumes. According to the United Nations Conference on Trade and Development (UNCTAD), maritime transport is the backbone of the global supply chain since the international shipping industry is responsible for the carriage of around 90% of world trade (UNCTAD, 2020). Shipping is a global service industry that, by general recognition, provides the lifeline of international trade. Suffice it to say that, due to the morphology of our planet, most international trade takes place by sea (Haralambides, 2019; Langen, 2020). However, essential to note that UNCTAD estimates that the operation of merchant ships itself contributes about US$380 billion in freight rates within the global economy, equivalent to about 5% of total world trade.

In this context, seaports are an essential spot in the global supply chain. Seaports are economic assets and value-adding logistics nodes with complex activity, comprising transactions related to cargo movement, whether in the transport or loading and unloading of vessels and which, if not responsibly

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

managed, can become a bottleneck in the supply chain. While cargo-handling operations form the core of ports, modern seaports have become logistics and industrial clusters, in addition to transport nodes (Langen, 2020; Parola, Satta, Notteboom, & Persico, 2020). As technological innovation, social transformation, and consumer behavior threaten firms' traditional boundaries, logistics need to adjust and adapt to new and emerging challenges, costs, and risks (Cacho, Marques, & Nascimento, 2020). Ports, however, are not an end but a means of passage. Being part of both more extensive transport and logistics supply chains and being a cluster of companies and businesses active in the sector, ports are in a unique position to fully grasp the potential generated by these new high-tech developments (Deloitte, 2021).

As long as humankind has been using water as a means of transportation, ports have served a critical purpose, permitting the transfer of goods and people from water to land. This unique role as a physical hub for freight flows has given ports a unique and longstanding role in information flow. In fact, Phoenician documents, dating from 1110 B.C., record trade transactions within the Phoenician network of Mediterranean ports (1). In many ways, this ancient network of ports coupled with a recordkeeping system could be considered an early form of geographically bounded information hubs in global supply chains (Srour, Oosterhout, Baalen, & Zuidwijk, 2008; Stopford, 2009). Ports are dynamic objects that constantly renew their technical systems according to the productive logic of each historical moment (Oliveira & Lima, 2016). Ports are growing in complexity. They are at the crossroads of many activities, modes, and stakeholders. They are essential multimodal nodes, hotspots for industrial activity, hubs for information, clusters for blue growth, etc. (Deloitte & ESPO, 2021). Modern maritime equipment constructors, ship owners and agents, transport and logistics companies, and port authorities collect enormous amounts of heterogeneous data at an unprecedented scale and pace (Herodotou, Aslam, Holm, & Theodossiou, 2020). Today port environments have become a locus of data generation and concentration. (Inkinen, Helminen, & Saarikoski, 2019). Technology is the most prominent force underlying change. Emerging trends – such as big data, data analytics, robotics, and artificial intelligence – show new means of supply chain coordination, other than the traditional hierarchical structure – in many instances, a chain of command – within the firm (Cacho et al., 2020). Areas of focus include establishing a vision for the future of maritime transport; identifying the innovative technologies, business models, and policies that will drive change; overcoming barriers to innovation; and establishing governance structures at the global and national levels to foster the innovations that our societies will need for a more sustainable and better performing future transport system (Rodrigue, 2010). Technology is the most prominent force underlying change. Ports dimensions have become intricate partner networks that include port authorities, terminals, shipping lines, trucking, logistics companies, and off-dock storage providers (Vanelslander & Sys, 2020). To be truly effective, stakeholders have to do more than adopt these technologies on their own. Instead, they must embrace platforms that make it easier for stakeholders to work together to promote the efficiency of the overall ecosystem (Delenclos, Rasmussen, & Riedl, 2018). "Platforms" are a term coined where many actors come together to share information and data on their activities. Within the port industry, 3 types of platforms can be distinguished (Deloitte & ESPO, 2021):

1. Port community systems
2. Cargo community systems
3. Terminal operating systems

Platforms increase efficiency by reducing paperwork and lowering the administrative burden. Nevertheless, producing and collecting data is not enough. In addition, by sharing real-time information, they

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

allow for increased transparency in the ecosystem and value chain, generating possibilities for just-in-time operations and optimizations (Deloitte & ESPO, 2021).

It is necessary to use them pragmatically to optimize flows, increase trade, and add economic value to related activities. For this, it is essential to connect, collect, monitor, monetize and optimize information (The Economist Intelligence Unit, 2020). Supply chain connectivity has the potential to become more robust. In this process, communication between different links in the supply chain is improved, processes are monitored more efficiently, and, consequently, operations are enhanced (Buxbaum, 2020). According to Ward and Bjørn-Andersen (2021), collecting and recording information that supports trade and the development of supply chains is not new: however, sharing that information is a much newer phenomenon (Ward & Bjørn-Andersen, 2021). Building a relationship with the maritime supply chain partners is imperative for organizations to survive and remain competitive (Osobajo, Koliousis, & McLaughlin, 2021).

Given their neutral position in both the supply chains and port ecosystem, ports are the perfect spider in the web to manage the digital oceans of platform data sharing (DNV GL, 2020). The amount of information generated by ports is constantly expanding as their applications. Data sharing between ports from different countries is still incipient and challenging (Lind, 2018). This challenge is also an opportunity for progress. Investing in digital infrastructure and new skills related to artificial intelligence, machine learning, and data science is a part of a new strategy in which new productivity levels can be achieved, contributing to improving efficiency at the various nodes along the maritime transport chain (Figure 1). Decision making in the industry can be improved by digital technology (Lind, Michaelides, Ward, & Watson, 2021)

Data use and sharing can engender significant opportunities to obtain new revenues, contribute to the identification of technological solutions to new or recurring problems and increase the competitiveness of the economic agents. With the generated knowledge, it becomes possible to identify choke points and business opportunities.

Creating value from data in maritime logistics depends on the data itself, the increasing maritime connectivity, the use of cloud computer algorithms, and the usage of machine learning. Data is the primary constraint to be faced: the data available must be sufficient in amount (quantity) and have a certain quality. Information will only create value if its amount is large and representative of the problem it is intended to solve (Kretschmann, Zacharias, Klöver, & Hensel, 2020).

Figure 1. The process of creating value from data

The use of machine learning in maritime logistics, in turn, comprises the advances in the Digital Supply Chain, the use of Digital Ships, and the improvement in Digital Ports (Kretschmann et al., 2020).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Supply chains are becoming more digital in recent years. The use of technological resources is being encouraged, and innovative ideas are being welcomed. Nowadays, it is possible to purchase raw materials, for example, without having to visit its producer, track the shipment throughout the supply chain, knowing in which container and in which vessel it is, its route, and its estimated time of arrival in each Port. Also, control and optimize stocks more efficiently through computer systems and benefit from more excellent connectivity between all stakeholders involved. This digitalization results in more transparency in the entire distribution network, taking full advantage of processes automation and cost reduction, providing better customer service, and providing digital data storage (Vieira & Fialho, 2020).

Shipping companies, in turn, are using more automated systems (Daniela, Calvert, Clark, & Parkhurst, 2019). Data is being collected during operations, and precise positioning systems are in use. These innovations help, for example, shipping companies optimize their fleet management through predictive maintenance since they have more information about cargo handling and fuel consumption. Data readings might suggest defects and alert the necessity of a certain kind of support. Besides that, routes can be better planned. Ultimately the decision-making process is optimized as well as the security at sea (T. E. Notteboom, Parola, Satta, & Pallis, 2017).

There are constraints to port management. The changing environment and the volatility in freight markets are stressing port operations, introducing structural changes in logistics and new distribution patterns. The large number of public and private stakeholders that interact in transport operations and the logistics of international trade generate coordination problems, requiring a considerable amount of information to be exchanged (Cacho et al., 2020)

A new logistics mantra could be spelled as "get the information flowing" (Cacho et al., 2020). Ports are only one node in the maritime supply chain, but an essential one (Brooks, Button, & Nijkamp, 2002; Carbone & De Martino, 2003). They must be able to adapt to the fast and changing conditions of international trade, ensure efficiency and safety in the handling and storage of cargo, enable the offer of value-added services, and be connected to other transport modes, such as highways, railways, and air transport in a cost-effective way.

Ports are evolving from traditional to digital ones. These last ones are converted into real technological parks and innovation environments (Heilig, Schwarze, & Voss, 2017). Nowadays, technology plays a significant role in them. The first step on becoming a digital hub is to promote intra-port coordination – through the adoption of computerized systems - and then build a connected network of ports within its country of belonging – through the systematic exchange of information. Ship to port and port to hinterland communication must be enhanced (Heilig, Lalla-Ruiz, & Voss, 2017; Kretschmann et al., 2020). Finally, connecting ports from different nations is an indispensable way to optimize and create value throughout the entire supply chain (Table 1).

The maritime industry is incrementally counting on the usage of data science. As impacts on the global supply chain, several benefits can be listed: new levels of efficiency, trade optimization, incident predictions, cost reductions, quality control improvement, reduced downtime and workloads, reliable forecasts of the movement of goods, simplification of information exchange, among others. Logistics then become more complex and dynamic.

Amidst this complexity, one cannot ignore the new realities that pressure economies like the pandemic or the growing consolidation of the Asian countries' productive sector. These changes in the contemporary world have repercussions on supply chains and logistics worldwide, motivated by people's heavy dependence on goods flows globally. Interruptions or changes in these flows significantly impact, demonstrating their vulnerability in an increasingly globally integrated world.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Usage of machine learning in maritime logistics – loci of application (adapted from Kretschmann et al., 2020)

	Digital Supply Chain	Digital Ships	Digital Ports
What it is	• The evolution of the supply chain, adapted to technological resources and innovative ideas from the manufacture of a product to its delivery to the final consumer.	• A set of various trends in the digitalization of shipping companies.	• The evolution of traditional ports, which are converted into technological parks and innovation environments.
What it comprises	• Raw material purchase. • Movement of goods. • Storage. •Transportation. • Other support activities.	• The use of precise positioning systems. • Systematic data collection during operations.	• Coordinated action between government, companies, and universities. • Intra port coordination. • Enhanced ship-to-port communication. • Improved port to hinterland connection. • Connected network of ports.
Its benefits	• Greater connectivity between stakeholders. • Process automation. • Optimization of stock control. • Digital data storage. • Cost reduction, • Better customer service.	• Better fleet management. • Optimization of processes, such as fleet maintenance, cargo handling, fuel consumption, and route planning. • Optimization of decision-making processes and security.	• The possibility of port data sharing. • New levels of productivity. • Time reduction of operations. • Gains in safety and environmental sustainability. • Automated and data-driven decision-making.

As a result of all of this, it is possible to observe significant changes in consumer behavior, translated into a reevaluation of production and logistic processes and accelerating digital transformation processes. These changes will allow a more intelligent communication in a global network concept; there sharing, integration, and management of data throughout the chain will be one of the determining factors of change (Carbone & De Martino, 2003; Vanelslander & Sys, 2020; Venkatesh, Zhang, Deakins, Mani, & Shi, 2020).

THE USE OF DATA SCIENCE IN BRAZILIAN AND PORTUGUESE PORTS

The relationship between Portugal and Brazil is unique thanks to deep historical and cultural affinities. Its economic dimension - one of the ten largest economies in the world - makes Brazil one of the most important countries for the internationalization of Portuguese companies, but the volume of trade and reciprocal investments, while significant, still has excellent potential for further growth. In addition to being bound to each other by a common language, the two countries have a privileged relationship since it is in their mutual interest to reinforce trade and investment flows. The relationship between Portugal and Brazil is framed by the Treaty of Friendship, Cooperation, and Consultation (*Brazil Ctry. Rev.*, 2021; Portuguese Government Portal, 2021). Geographical distance plays a much lesser role today as a determinant of trade between countries, and it is being replaced in trade models by the concept of economic distance, as this is proxied by ocean freight rates (cf. transport costs). Ports today, technical-scientific-informational period, objects are essential to the smooth flow of goods across several territories and thus fundamental to the business enterprises and states. To meet the requirements of the globalized world

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

and the growing movement of goods in their docks, ports are modernized. In turn, the advancement of technical systems in ports contributed to greater flexibility in the movement of goods and the reduction of waiting time of cargo and ships on the docks (Oliveira & Lima, 2016). These developments have expanded the international markets and have improved the international competitiveness of exporting countries, and it has facilitated the industrialization of many of them around the world (Haralambides, 2019). Connected technologies are moving ports into the digital age and helping to boost productivity and cut costs (Delenclos et al., 2018). Digital and physical connectivity go hand in hand (UNCTAD, 2019).

The Brazilian Context

According to the Brazilian Infrastructure Ministry Homepage (Ministério da Infraestrutura, 2020), there are 36 public seaports and 39 public river ports in Brazil. Besides them, there are approximately 128 private use terminals. These last ones are port facilities operated under the government's authorization and located outside the public port area.

From north to south, the situation from these different ports varies a lot: there is no homogeneity between them. On the one hand, there are harbors devoid from the use of technology where humans carry out many manual steps and where data is only available in sparse documents. On the other hand, there are public ports and private use terminals that obtain productivity gains each year due to the massive use of technology and improve efficiency, speed, and accuracy. The reality along the coastline is quite complex (Mendes Constante, Langen, Vieira, Lunkes, & van der Lugt, 2018).

However, it is possible to state that all terminals that carry containers – regardless of whether they are public or private – have certain harmony among Brazilian ports. They are all automated and highly efficient. All of them invested, in recent years, in digital infrastructure, news skills and obtained new levels of productivity. In general, one may say that the treatment of containers is different from that of general cargo since the first ones are responsible for moving goods with higher added value. Therefore, Brazilian container terminals have stricter control over their cargo (Vedan & Thomas, 2019).

Because of this disparity between ports that handle containers and ports that handle other types of cargo, the existing data is not always good in terms of quantity or quality. Thus, it is hard to talk about an overall system that coordinates information between all the existing ports in Brazil, a country whose coastline is 7,491 kilometers long. Intra port coordination is still lacking (Rodrigues, Ferreira, Murta, & Murta, 2017).

In order to fill the present gap, there are initiatives promoted by the Federal Government that aims to collect data on port operations and their respective prices and generate information from them. One of these measures is promoted by the National Waterway Transportation Agency (ANTAQ - *Agência Nacional de Transportes Aquaviários*) - a Regulatory Agency - and consists of a system called "Port Performance System" which is considered as a strategic tool for planning and a significant data source for the generation of statistics and sector indicators (OECD, 2019). More than collecting data, the system aims to produce information that serves as a tool for a range of activities (Table 2).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. The purpose of the "Port Performance System"(adapted from ANTAQ, 2019)

Generate information that serves as tools for:
• Operational management at the terminals of each public Port or private use terminals, detecting possible deficiencies and dysfunctions and enabling diagnostic and correction actions. • Port development planning by comparing port cargo handling with capacity data and facility and equipment utilization levels, detecting trends and needs to promote expansions and improvements. • The monitoring of results derived from strategic measures and actions adopted by the Federal Government or by Port and terminal managers and operators. • Public knowledge and control of the performance and tariffs of port services offered, allowing a complete assessment of their quality and reflections on total transport cost. • Regulation of the economic activity. • The achievement of standards and comparative parameters of performance and tariffs between various facilities and terminals.

In this system, public ports and private use terminals provide data upon:

1. **Mooring:** place, International Maritime Organization (IMO) number, shipowner's nationality, operation type, etc.
2. **Temporal Data:** date/time of vessel's arrival, estimated date/time of mooring, date/time of mooring, date/time of unberthing, date/time of start/end of the operation.
3. **Stoppages:** date/time of start/end, reason (rain, anchoring without forecast of mooring, lack of electricity, strike, port equipment break, accident, tide, and others).
4. **Monetary Data:** amount paid for the use of waterway infrastructure, terrestrial infrastructure, mooring fare, equipment fare, the amount paid for internal transportation and labor, other rates.
5. **Cargo:** company responsible for the port operation, type of cargo (solid or liquid bulk, general cargo), full or empty (if containers), Bill of Lading number, the direction of cargo movement (loading or unloading), country of origin/destination, quantity, gross load weight.
6. **Types of Cargo Operation:** support, supply, long-distance navigation (between ports from different countries), short sea shipping (between ports in the same country), cargo transfer, etc.
7. **Containerized Cargo:** cargo code (related to the cargo moved inside the container), Bill of Lading number, net cargo weight, and container identification (alphanumeric field with 11 digits).

Another important initiative promoted by the Federal Government is the "Paperless Port" (PsP – *Porto sem Papel*) system, which automated the provision of information necessary for the stay of vessels in Brazilian public ports. The PsP aimed at improving port efficiency by reducing bureaucracy and the volume of redundant activities performed by agents involved in releasing, loading, and unloading operations (Governo do Brasil, 2021; Ministério da Infraestrutura, 2018).

Its main objective is to gather data and documentation to speed up the analysis and release of goods. Many different intervenient authorities are involved in using the system, such as the Port Authority itself, the National Health Surveillance Agency, the Maritime Police, the Ministry of Agriculture, Livestock and Supply, Brazil's Navy, and the Customs Authority. Among the advantages of its usage are standardization of information, reduction of bureaucracy, rationalization of procedures, agility, predictability, and transparency (Fioresi & Guimarães, 2016; Rodrigues et al., 2017).

As said before, a harbor is an important node in the supply chain and a great data generator. Although historically, traditional ports have been making decisions based on experience and the country's economic context, the value of data-based decision making is unquestionably and an essential component

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to improving performance. Data-driven decision-making assumes the function of combining, analyzing, and visualizing data from different sources and providing it as an input for decision-making processes (Kretschmann et al., 2020).

There is, although, a considerable gap between theory and practice, especially in Brazil. With its vast coastline, there is a myriad of different economic agents involved in port activities. Public ports and private terminals and ports that handle containers and ports that handle other types of cargo present vastly different dynamics along the shoreline. It is not easy to homogenize and consolidate the distinct data provided (Vieira & Fialho, 2020).

If intra-port coordination is still lacking, there is a long way to go in order to promote a connected network of ports with a different country.

The Portuguese Context

Portugal has a long maritime tradition, and although its coastline is not as long as the Brazilian one, it is very well developed. Portugal has 14 continental ports divided between the ones dedicated to cargo handling and the ones dedicated to fishing and recreation. Among them, the Ports of Sines, Viana, Leixões, Aveiro, Figueira da Foz, Lisbon, Portimão and Setúbal are prominent. Besides these, there are important ports in the Azores Archipelago and the Madeira Archipelago.

In 2017, through Resolution of the Council of Ministers no. 175/2017, a national strategy was approved in order to increase the competitiveness of the continent's commercial ports network. With this program, the country intends to invest in the maritime economy, targeting boosting economic activities at sea, creating business opportunities that lead to job creation, and increased exports, maximizing the dynamics of maritime transport. This strategic vision is based on three pillars: i) Portugal as a global logistics platform that generates value; ii) creation of a business accelerating port hub capable of attracting investment and supporting the internationalization of the Portuguese economy; and iii) Portugal as a hub for liquefied natural gas (LNG) in the Atlantic (DRE, 2017).

In order to achieve this strategic vision, each Portuguese Port must combine data from different sources, systematically forecasting events or potential outcomes with predictive analysis, providing a new level of port management. Ports ought to become more digital and sustainable, abandoning their traditional profile. There is a large amount of data produced by Portuguese ports, and they can be divided into the following categories:

1. **Port System Assets:** comprising infrastructure, equipment, vehicles, buildings, vessels, communication networks, hardware, software, and other applications. These data characterize the assets that make up a port system.
2. **Cargo Operations:** data concerning the movement of goods and the information necessary to plan and control operations at the Port and between ports.
3. **Vessel Operations:** data concerning the information necessary to plan and control ship handling operations in ports.
4. **Access:** data related to the control and flow of people and goods in the port area.
5. **Multiple Statements:** general data corresponding to the necessary formalities for the entry and departure of ships and goods from the Port.
6. Maritime Traffic at and near the Port.
7. Meteorology, Waves, and Tides.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

8. Energy Consumption and other utilities.

Table 3. Axes of action in the promotion of Digital Ports (adapted from Miranda, 2020)

	Axis 1	**Axis 2**	**Axis 3**
What it is	• Digitalization	• Decarbonization	• Accessibility
What it comprises	• The fifth generation of mobile communications (5G) applied to ports. • Better and faster connectivity.	• The replacement of traditional propulsion fuel with alternative energies.	• The strategy of gaining scale to receive larger ships and optimize the flow downstream.
Its application	• In the monitoring processes of containers and goods. • In loading and unloading operations.	• In ships. • In ports. • In general operations.	• In the navigation channels. • At terminals and road or rail access.
Its benefits	• Integrated information processing. • Efficiency gains in port processing. • The workforce may be progressively supplemented by port computer operators. • Stricter enforcement.	• The storage and processing of the generated waste will improve.	• Elimination of bottlenecks. •Better intermodality. • Simplification of the entire supply chain.

European Commission considers a need to promote high-performance seaport activities through strategic actions to connect seaports to the trans-European network, modernize port services, attract investment, promote social dialogue, improve the environmental profile, and encourage innovation. Furthermore, through the Directive 2010/65/EU, the EU aims to increase the opportunity to facilitate trade and administrative barriers through harmonization and standardization of the European Seaports Organization. Entities of the State sphere carry out the implementation of the Directive. Member States show a low priority to exchanging information between the Member States, and the majority has not taken any action in this direction. Concerning the Single Window Environment development model, the mixed model is adopted in most Member States (Jose & Fernandes, 2020; Linke, 2012).

The term Single Window Environment was introduced by World Customs Organization (WCO). Based on the definition presented in Recommendation No. 33, and is widely used in the area of international trade as "an 'intelligent,' facility that allows parties involved in trade and transport to lodge standardized information, mainly electronic, with a single entry point to fulfill all import, export, and transit-related regulatory requirements" (WCO, 2021).

It was decided that in Portugal, the Logistics Single Window (JUL - *Janela Única Logística*) would be developed considering all the necessary legal and operational requirements according to the EMSWe (European Maritime Single Window Environment) concept. The JUL is currently being developed by the Ministry of the Sea, through the Directorate-General for Natural Resources, Safety and Maritime Services (DGRM) and all the Portuguese Ports, together with a broad group of partners (Projeto JUL, 2021).

JUL is a digital ecosystem that allows data sharing between port authorities and other national authorities and organizations at the level of the European Commission. Port cooperation for the Portuguese is not something new. The platform is already contributing to making Portuguese ports less bureaucratic and paperless, with noticeable productivity gains. Moreover, as information is shared among different agents, a degree of data homogenization is achieved. Besides that, associated with the technological platform, a legal framework guarantees consistency, management, and operation in a coordinated and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

harmonized way. It is possible to identify that intra-port coordination in Portugal and a connected network of European ports already exist.

Nowadays, the focus is on endow the country with "new generation ports" or Digital Ports, fully automated and environmentally friendly. According to Miranda (2020), to achieve so, actions must be focused on three axes (Table 3).

Portuguese ports are investing in digitalization, decarbonization, and accessibility (Portugal.EU, 2021). Through the first one, new investments are being made in telecommunication to achieve better and faster connectivity. The results are integrated informational processes, gains in efficiency, progressive replacement of human labor by port computer operators, and the possibility of carrying out more detailed and rigorous inspections (Czermański, Cirella, Oniszczuk-Jastrząbek, Pawłowska, & Notteboom, 2021; Notteboom, Lugt, Saase, Sel, & Neyens, 2020). Through the second one, Portugal is aligning itself with the trend of making ports environmentally friendly. This step is an important one to promote Portugal as a global logistics platform that generates value. The accessibility axis is intended to work on the existing supply chain's chokepoints, facilitate a better intermodality, and boost the simplification of the entire supply chain (DG MOVE, 2019).

The country wants to shift its port model from a traditional to a digital one and transform them into genuinely international ports. For this to happen, it is necessary to invest in technological infrastructure and new partnerships. In this context, it makes sense to search for partnerships with countries like Brazil, with which it shares a common history (in addition to the Portuguese language), and in which the largest Port in Latin America is located.

Other broader initiatives have already taken place, such as the "Portuguese Language Ports Association," which aims to ensure the defense and promotion of its members' interests and contribute to the port development and modernization. Its members are Angola, Brazil, Cape Verde, Saint Thomas and Prince, Guinea-Bissau, Equatorial Guinea, Mozambique, Portugal, East Timor, and Macau (China).

In this context, data is a strategic asset for the Portuguese ports. On the one hand, they constitute the basis of the information; on the other, the information constitutes knowledge. Its use in a broader way is already frequent in the country.

Given its unique nature, the data has assumed a predominant role in preparing information to support the decision and improve business processes. Many authorities take part in the process of data analysis and generation of information. Some of the information produced by ports is submitted to official entities in compliance with obligations. The National Institute of Statistics stands out, and the Budget Support Technical Unit, the Mobility and Transport Authority, the Transport Mobility Institute, and the General Direction of Natural Resources.

The Possibility of Port Data Integration

The world in which ports are tasked to facilitate global transport flows and create stable, resilient port infrastructure has become far more complex over the past decades. This has not lessened port authorities' core role and function but only added to their complexity. Port authorities are still neutral players in an increasingly competitive market environment, improving connectivity, facilitating private companies, attracting investments. However, they are also developing new roles to serve customers better, increase added value to customers, and meet general interest obligations. These evolved and new roles are sometimes a direct result of increased demand for public tasks, such as more focus on sustainability or a need for a neutral innovation facilitator. In other cases, they have become a business opportunity (although never

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in competition with the port clients), for example, through investment in renewable energy generation or the development- and sale of in-house digital tools. (Deloitte & ESPO, 2021).

Traceability is at the core of modern supply chain management. For this to happen, information must be easily shared and freely exchanged amongst participants. The premise underneath the idea of port data integration is the hypothesis that a large amount of data is available. In a dynamic environment, intense IT technologies, standards, and harmonization processes are essential. In turn, better supply chain coordination requires further integration. This reorganization has substantial implications for management from risk exposure and cost efficiency (Cacho et al., 2020).

A multidimensional conceptual framework for port logistics integration is proposed to incorporate the role of the three infrastructural variables emerging from the recent developments in the port logistics environment (Grifoll, Ortego, & Egozcue, 2019). The literature review has found the logistics process and operations, information integration, value-added services, logistics practices, organizational activities, resource sharing, and institutional support as influential factors in logistics integration (Alavi, Sayareh, Fei, & Nguyen 2018).

In the case of Brazil, it was possible to detect that although ports are a spot in the supply chain that engenders lots of information, not all ports in Brazil are at the same pace. With a vast coastline, there are disparities between harbors, but there are initiatives from the Federal Government to homogenize and centralize information. It is crucial to notice that the process of port modernization in Brazil started to took place in 1993 with the new legislation on ports and, since then, the different harbors have been modernizing at a special rate (Rodrigues et al., 2017).

Portugal, in turn, is a much smaller country with a tremendous and traditional vocation towards the sea. The different continental ports already share information at the same time as they share information with European Authorities. The experience with cooperation seems to be more sophisticated. Meanwhile, Portugal considers South America a strategic region for its business and hopes to increase maritime trade with Brazil.

In order to achieve the intended port integration between countries, homework must be done before: the existing intra-port cooperation must be more elaborated. If, in theory, it all may seem simple and obvious, there are many to overcome in practice. One obstacle seems to be the absence of homogeneity between Brazilian ports, as said before. Another one is that many data provided by Brazilian ports to authorities are protected under business confidentiality. It means that not all information generated is available and can be shared.

Despite the constraints, Portugal can become Brazil its primary gateway to Europe. Located at the confluence of important international maritime routes, taking advantage of its Euro-Atlantic centrality, Portugal emerges as a hub for the European market with developed networks of logistics and transport infrastructures. It is also an alternative to other congested European ports. The European country can also contribute to Brazilian ports with its experience in terms of innovation.

Brazil, in return, is a much bigger economy, and this gap among countries can be seen as an opportunity (not as another obstacle) that can generate gains for both nations. Table 4 demonstrates the mismatch between the two economies.

According to data from UNCTAD, Brazilian ports process almost four times more containers than Portuguese ports (1.36% of world share against 0.36% in 2019). The Port of Santos, for instance, is the largest in Latin America and is the most important foreign trade route in Brazil. The South American country is also responsible for a much larger share of world exports (11.19% against 0.35% in 2019). There is still a disparity among economies in terms of merchandise imports, but not as big as the ones

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Comparison between Brazilian and Portuguese maritime economy in 2019 (adapted from UNCTAD, 2021a, 2021b)

	Brazil	Portugal
Container port throughput	10,982,131.00 TEU 1.36% of the world share	2,920,700 TEU 0.36% of the world share
Number of port calls	29,807.00	13,687.00
Merchandise exports in US$, in the world share	11.19%	0.35%
Merchandise imports in US$, in the world share	0.96%	0.47%
Top 5 partners	China United States of America Netherlands Argentina Japan	Spain France Germany United Kingdom United States of America

mentioned before. Brazil's economic prominence may be beneficial to the Portuguese and may also be an essential gateway to the regional market. A barrier that must be surpassed is the low connectivity between the two countries. In Europe, the Netherlands is the major Brazilian economic partner. Nevertheless, before it, China and the United States of America play a significant role. Portugal, in turn, has its most important economic partners centered in Europe (UNCTAD, 2021a).

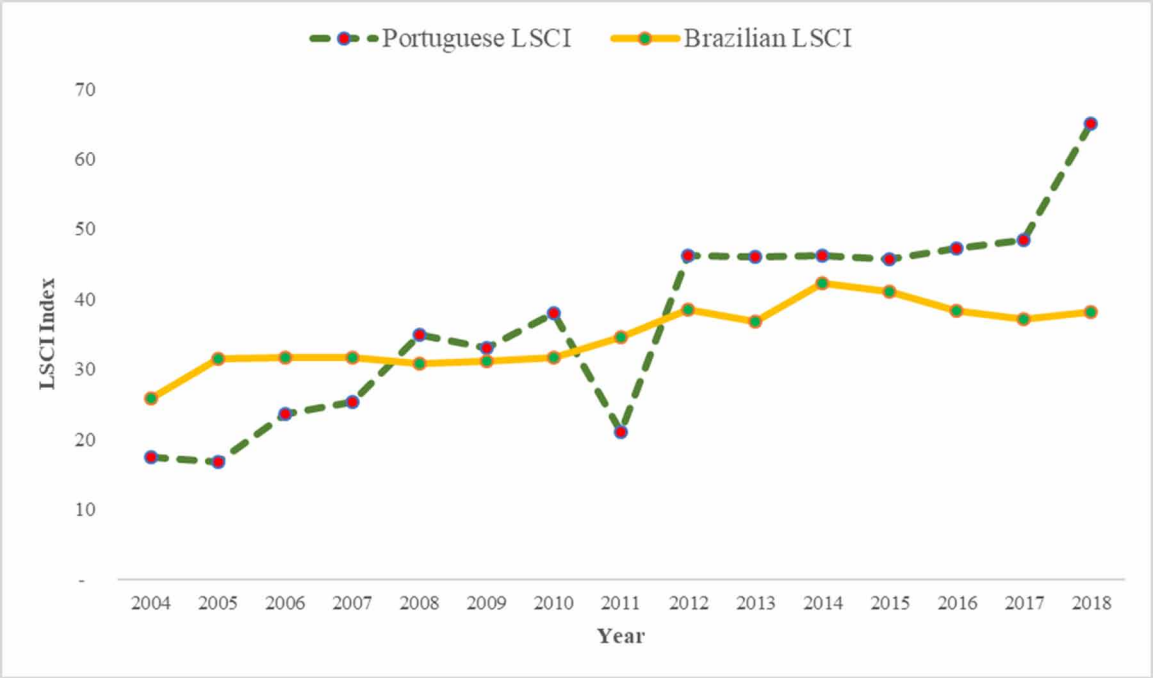
Efficient and well-connected container ports enabled by frequent and regular shipping services are crucial to minimizing trade costs, including transport costs, linking supply chains, and supporting international trade. Thus, port performance is a critical factor that can shape countries' trade competitiveness. Every hour of port time saved by ships translates into savings in port infrastructure expenditure for ports, ship capital costs for carriers, and inventory holding outlays for shippers (World Shipping Council, 2021).

Recognizing the importance of measuring container port performance, UNCTAD developed the Liner Shipping Connectivity Index (LSCI) in 2004 to determine countries' positions within global liner shipping networks; the latest country-level LSCI statistics were published in July 2019. LSCI is generated from five components: (a) the number of ships; (b) the total annual container-carrying capacity of those ships; (c) the maximum vessel size; (d) the number of services; and (e) the number of companies that deploy container ships on services from and to a country's ports. The data are derived from Containerisation International Online (until 2015) and MDS Transmodal (from 2016 onwards). The index is set at 100 to the maximum value of the country's connectivity (UNCTAD, 2020).

Brazil's connectivity index has been constant since 2006. When it was created, as shown in the chart below, but always between 30 and 40 points. Their top ports are Port of Santos, Rio Grande, Rio de Janeiro, and Paranaguá. The Portuguese index had grown substantially since 2006, starting with a lower score than the Brazilian one, reaching 60 points in 2018, when it dropped due to the international economic crisis (Figure 2). Even after the crisis, Portuguese ports' connectivity is higher than Brazilian ports' connectivity. The top ports in Portugal are the Port of Sines, Leixões, Lisbon, and Setúbal. Once again, the inequality between the two countries takes place (UNCTAD, 2021a, 2021b).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Brazilian and Portuguese LSCI (adapted from UNCTAD, 2021a, 2021b)



It seems that a vast program of port cooperation and data integration among Brazil and Portugal is still distant from happening since constraints were observed as summarized below (Figure 3). Even inside Brazil, there is no homogeneity of ports in the usage of digital infrastructure, cloud computing, or artificial intelligence. On the one hand, the existing inequalities hinder general cooperation between nations but, at the same time, reveal opportunities to approach specific nodes in the international supply chain. The research purpose is to start the debate about port data integration, its opportunities for optimization and value creation among these two Portuguese-speaking countries. Is it so straightforward to achieve it, or the reality poses obstacles to its realization? What can be accomplished at the current stage of data science application?

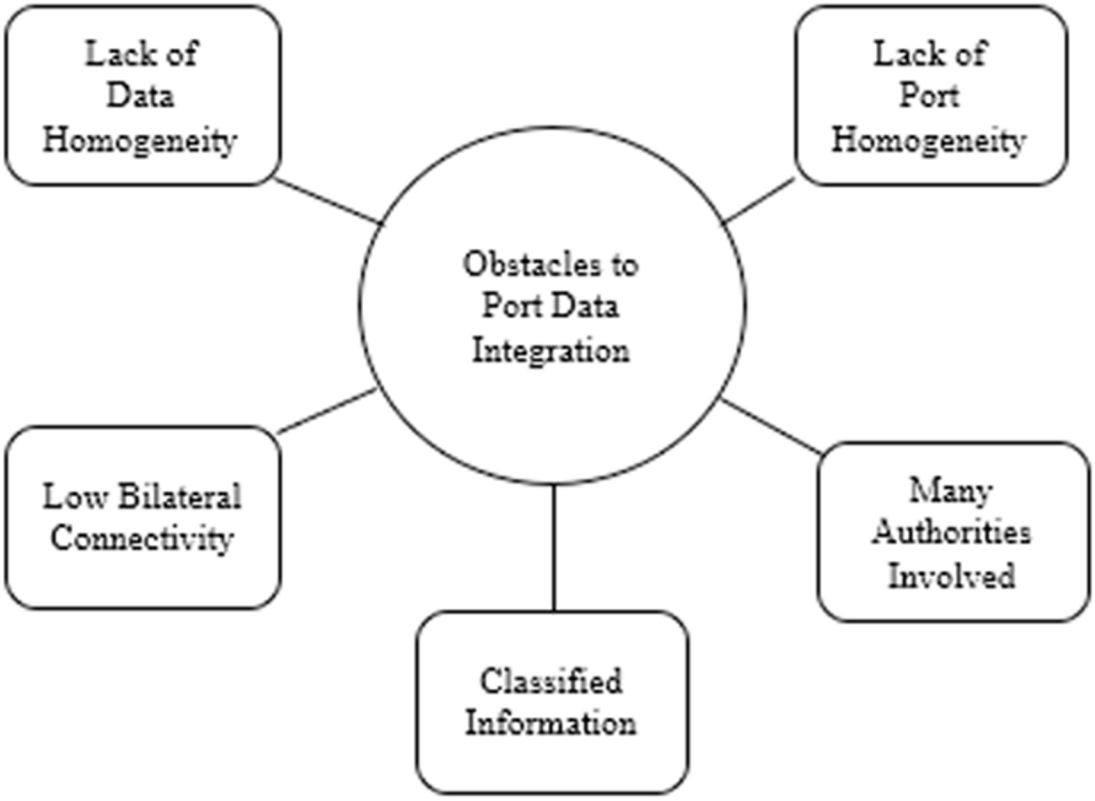
Suppose, at first sight, port data integration is seen as a challenge. In that case, it must also be seen as a business opportunity for new revenues, identifying technological solutions to new or recurring problems, and increasing economic agents' competitiveness (McKinsey, 2020).

A start on the way to overcome these difficulties is to start with a less bold project. One thing that can be done is to identify in each country its potentialities. It means: determine which particular Port can start the process. For example, Port of Sines in Portugal, for example, according to the LSCI, seems to be way ahead, with 45,20 points, from Port of Leixões (16,27 points) and Lisbon (16,12 points). In the Brazilian case, ports that handle containers, as Port of Santos (35,79 points), Paranaguá (32,22 points), Rio Grande (30,55 points), and Rio de Janeiro (29,16 points) emerge as promising candidates for a cooperation program.

What can be done to improve a port's connectivity? According to UNCTAD (2021b), "go digital" is essential to enhancing port connectivity since digital and physical connectivity go hand in hand. Just as trade benefits from the latest technologies such as artificial intelligence, the Internet of Things, and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Obstacles on Port Data Integration



blockchain, port and shipping operations would also benefit from taping the opportunities arising from digitalization (UNCTAD, 2021b).

After determining which ports in each country are data-rich and more connected with the international port network, it is possible to draw the desired outcome from the data integration process, frequently characterized as optimization in cost-efficiency, reduced response time, and increased safety. Data interchanging will provide digital intelligence storage for agents involved. From that point, they will detect trends, generate knowledge and identify chokepoints and opportunities.

One example can be sketched. Within the logistics industry, cargo community systems have been quite prominent in the past decade. A lack of relevant and accurate information in the supply of data can result in inefficient processes and other 'hidden costs' such as excessive inventories. By sharing more information, it is possible to optimize these processes and reduce costs. Fast and reliable cargo information systems have become increasingly indispensable for international transport hubs (Deloitte & ESPO, 2021). Imagine if Port of Sines and Port of Paranaguá sign a cooperation term to exchange information about cargo transacted between the two ports. The main objective of this cooperation is to guarantee a better integration among them, increase flow transparency, and ultimately reduce operating costs and time and increase operation security. It is well known that containers are scanned when exported and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

imported. Port of Paranaguá inspects 100% of cargo that arrives at the Paranaguá Container Terminal (TCP) and enters the Port's public area.

The scanning operation brings safety to the process, but it requires time and money. The scanned images are sent directly to the Customs Service system, which crosses, in real-time, the captured image with data described in the invoice. In this way, the inspection process is faster and more accurate. The scanner can also detect any type of illicit or smuggled substance, such as chemicals, weapons, drugs, and radioactivity.

If there is an agreement that defines that the container must be scanned only in the Port of origin or destiny with image sharing between ports that subscribed to the cooperation agreement, there would be no need for a redundant operation in two points of the supply chain that costs time and other resources. Besides that, there can be gains in safety since there would be more time and scanners available to "read" other cargo than the same one for the second time.

Although difficult at first sight, port data integration opens opportunities for optimization (like the one described above) and value creation. However, it is vital to have in mind that the steps to be taken must be small and punctual at first.

SOLUTIONS AND RECOMMENDATIONS

The competitiveness of a seaport highly depends on its efficiency, especially in terms of logistics practices, functions, and activities and how all that is integrated with those of the other players in the supply chain (Alavi et al., 2018).

If, on the one hand, the Academy has been focusing on issues concerning data science, artificial intelligence, and cloud computing and there is a wide range of scientific articles and literature on these topics, on the other hand, it is possible to observe that there is still much to be done in this field in the port sector. There are successful experiences in the use of new information technology at specific points in the supply chain. Port of Sines in Portugal and container ports in Brazil are examples of supply chain nodes that use data science and can be considered true Digital Ports. Nevertheless, port data integration between ports from different countries is a more complex process.

Projects that aim for data integration should start more humbly. Instead of thinking big and try to connect two entirely different realities, it is essential to detect which actors can effectively make a difference and contribute to the progress. It is crucial to make an accurate diagnosis and start with a small-scale program. There will be no generic solutions, but each data exchange and integration process must be seen as unique and will face its obstacles and successes.

Furthermore, it is essential to coordinate action between government, ports, and universities to produce possible cooperation arrangements and find creative solutions.

FUTURE RESEARCH DIRECTIONS

Ports are growing in complexity. They are at the crossroads of many activities, modes, and stakeholders. Digitization allows them to cooperate and integrate along the supply chain and optimize, manage, and automate processes. Ports are actively becoming a digital hub, with port authorities as spiders in the web in which all stakeholders are connected and jointly collaborate. This is a central role for port authorities

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

as neutral players to facilitate data exchange between more commercial players. Dissemination of the effect of port activities and the sharing of data and insights will increase the engagement of the surrounding communities (including the academic and business community) and generate a competitive edge (Deloitte & ESPO, 2021).

The Academic community has already started studies on the applicability of data science, machine learning, cloud computing, and artificial intelligence in the port sector. Some theoretical and generic studies on the subject have already been published lately. What is needed is to start carrying out more practical studies, such as case studies. It is necessary to analyze in the specific case how this application happens.

Initially, it is necessary to carry out case studies of ports at the forefront of this process, such as the Port of Sines in Portugal. This type of empirical research must then be extended to bring state-of-the-art knowledge to different ports, such as those located in Brazil. What initiatives have these ports taken to become Digital Ports? How do they contribute to making digital supply chains more digital? What has been done to promote more excellent connectivity of these ports? These researches are yet to be done.

Only after getting to know the local realities can we outline models of cooperation. These models will be helpful in the sense that they will identify which results can be expected, achieved, and measured. In other words: what can be done in practice and what belongs to the world of theory. How to measure the gains on security, time, cost, and productivity derived from this integration process? What are the indicators to be used? They still need to be modeled. This is also a research direction on the topic.

CONCLUSION

The greening and digital transitions combined with the changing global political environment and the impacts of the pandemic create the perfect storm for the port industry. Ports are at the crossroads of these transitions and changing realities. The diverse challenges ahead, the increasing complexity and scale in the sector, push ports to cooperate with other ports: from coalitions on a single project to complete mergers. By teaming up, ports can leverage external knowledge or lower the risk of particular investments (Deloitte & ESPO, 2021).

The initial debates among researchers regarding the state of the art in applying data science in the port sector were marked by enthusiasm regarding its possibilities. When talking about technology, a new world unfolds. The reality is that different actors are at different stages of development concerning this subject: development is uneven. This inequality creates obstacles in cooperation, but it also generates opportunities, not only for business but also for learning.

An observer may think that Brazil and Portugal can quickly enter into cooperation agreements to exchange data and information since both countries speak the same language (Portuguese) and have a centuries-old connection. Nevertheless, when looking at a specific economic sector closely, such as the port sector, inequalities become apparent.

Initially, the coast size of both countries is too divergent: Brazilian coastline is 7,491 kilometers long, one of the longest in the world, while the Portuguese coastline is 943 kilometers long. Therefore, the Brazilian coast has many more ports than the Portuguese one. The rich Brazilian diversity in data homogenization and the construction of an efficient communication system between ports can hinder the process. Despite the size of the national port sector, it was observed that those ports that handle containerized cargo have massively invested in technological infrastructures, gained productivity, and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

adopted new strategies for insertion in the global maritime network. These agents constitute a possibility for cooperation (Haralambides, 2019).

On the other hand, Portugal is a country with a small coastline, maritime tradition, and a strategic position, becoming a bridge between the Atlantic, the European continent, and the African continent. Their reduced number of ports compared to Brazil has made cooperation possible between them and other supranational authorities, such as the European Commission. The cooperative processes of which the country is a part already have a more sophisticated framework.

After the initial enthusiasm about the myriad of possibilities for applying data science in the port sector, it was observed that the development of the subject is uneven among the countries under study and within each of those countries (especially in the case of Brazil). It was concluded that a broad project covering all national ports would not be possible at the moment. It was necessary to realign expectations.

Any port data integration project will necessarily involve identifying those agents who can contribute to the process. In Brazil, for instance, it was observed that not all ports would be ready for this. After this delimitation, it is necessary to determine what is expected from this integration. Security gains? Gain in operational time? New recipes? From this specific cooperation, the process can expand and begin to face the identified obstacles: lack of data homogeneity, lack of port homogeneity, the fact that there are many authorities involved and it is necessary to find a way to cope with them, the use of classified information and the low bilateral connectivity. PsP and JUL are a good base for a start.

Portugal and Brazil have a privileged relationship. Considering that digital and physical connectivity go hand in hand, the process of port data integration has to be considered between the two countries since it is in their mutual interest to reinforce trade and investment flows. Just as trade benefits from the latest technologies such as artificial intelligence, the Internet of Things and blockchain, port and shipping operations would also benefit from taping the opportunities arising from digitalization (UNCTAD, 2019). Port data integration helps leverage external knowledge, engage stakeholders, create new decision-making anchors, lower the risk of certain investments, boost productivity and cut costs, accelerate greening and digital transition, and generate possibilities for just-in-time operations and optimizations.

REFERENCES

Alavi, A., Sayareh, J., Fei, J., & Nguyen, H.-O. (2018). Port Logistics Integration: Challenges and Approaches. *International Journal of Supply Chain Management, 7*.

ANTAQ. (2019). *Port perfomance system: external user manual*. Agência Nacional de Transportes Aquaviários – ANTAQ.

Brazil Country Review. (2021). Retrieved from https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=149896693&site=eds-live

Brooks, M. R., Button, K., & Nijkamp, P. (2002). *Maritime Transport*. Edward Elgar Pub.

Buxbaum, P. (2020). Pandemic developments will drive future supply-chain improvements. *American Journal of Transportation*.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cacho, J. L., Marques, L., & Nascimento, Á. (2020). Customer-Oriented Global Supply Chains: Port Logistics in the Era of Globalization and Digitization. In V. Chkoniya, A. O. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 82–103). IGI Global. doi:10.4018/978-1-7998-3115-0.ch005

Carbone, V., & De Martino, M. (2003). The changing role of ports in supply-chain management: An empirical analysis. *Maritime Policy & Management*, *30*(4), 305–320. doi:10.1080/0308883032000145618

Czermański, E., Cirella, G. T., Oniszczuk-Jastrząbek, A., Pawłowska, B., & Notteboom, T. (2021). An Energy Consumption Approach to Estimate Air Emission Reductions in Container Shipping. *Energies*, *14*(278), 278. doi:10.3390/en14020278

Daniela, P., Calvert, T., Clark, B., & Parkhurst, G. (2019). *New Technology and Automation in Freight Transport and Handling Systems*. Academic Press.

de Langen, P. W. (2020). *Towards a Better Port Industry: Port Development, Management and Policy*. Taylor & Francis. doi:10.4324/9780203797501

de Miranda, A. S. (2020). *Uma nova dinâmica para os portos portugueses*. PÚBLICO.

de Oliveira, W. B., & Lima, L. C. (2016). Ports and flows in the technical-scientific-informational milieu: The case of the Port of Fortaleza. *Caderno de Geografia*, *26*(46), 597–614. doi:10.5752/p.2318-2962.2016v26n.46p.597

Delenclos, F.-X., Rasmussen, A., & Riedl, J. (2018). *To Get Smart, Ports Go Digital*. BCG.

Deloitte. (2021). *Smart Ports - Point of View*. Academic Press.

Deloitte & ESPO. (2021). *Europe's ports at the crossroads of transitions A Deloitte and ESPO Study.* European Sea Ports Organisation (ESPO) and Deloitte Port Advisory.

DG MOVE. (2019). Mobility and Transport Transport in the European Union Current Trends and Issues Background Information. Brussels: Directorate-General for Mobility and Transport (DG MOVE) of the European Commission.

DNV GL. (2020). *Ports: green gateways to Europe*. DNV.

DRE. (2017). Resolução do Conselho de Ministros 175/2017, 2017-11-24 - DRE. *Diário Da República n.º 227/2017, Série I de 2017-11-24*. Retrieved from https://dre.pt/home/-/dre/114248655/details/maximized

Fioresi, E., & Guimarães, L. (2016). *Brazilian management port: a socio-technical approach*. Academic Press.

Governo do Brasil. (2021). *PSP - Porto Sem Papel*. Retrieved from https://concentrador.portosempapel.gov.br/PSP-CDP/private/comum/pages/home.xhtml

Grifoll, M., Ortego, M. I., & Egozcue, J. J. (2019). Compositional data techniques for the analysis of the container traffic share in a multi-port region. *European Transport Research Review*, *11*(1), 1. doi:10.118612544-019-0350-z

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Haralambides, H. E. (2019). Gigantism in container shipping, ports and global logistics: A time-lapse into the future. *Maritime Economics & Logistics*, *21*(1), 1–60. doi:10.105741278-018-00116-0

Heilig, L., Lalla-Ruiz, E., & Voss, S. (2017). Digital transformation in maritime ports: Analysis and a game-theoretic framework. *NETNOMICS: Economic Research and Electronic Networking*, *18*(2-3), 227–254. Advance online publication. doi:10.100711066-017-9122-x

Heilig, L., Schwarze, S., & Voss, S. (2017). *An Analysis of Digital Transformation in the History and Future of Modern Ports*. doi:10.24251/HICSS.2017.160

Herodotou, H., Aslam, S., Holm, H., & Theodossiou, S. (2020). Big Maritime. *Data Management*, 313–334. doi:10.1007/978-3-030-50892-0_19

Inkinen, T., Helminen, R., & Saarikoski, J. (2019). Port Digitalization with Open Data: Challenges, Opportunities, and Integrations. *Journal of Open Innovation*, *5*(2), 30. doi:10.3390/joitmc5020030

Jose, P. C., & Fernandes, A. P. (2020). European Seaports Information Systems. The Impacts of Directive 2010/65/EU. *Economics and Culture*, *17*(2), 38–49. doi:10.2478/jec-2020-0019

Kretschmann, L., Zacharias, M., Klöver, S., & Hensel, T. (2020). *Machine Learning in Maritime Logistics*. Fraunhofer CML.

Lind, M. (2018). Digital Data Sharing: The Ignored Opportunity for Making Global Maritime Transport Chains More Efficient | UNCTAD. *UNCTAD Transport and Trade Facilitation Newsletter, 79*.

Lind, M., Michaelides, M., Ward, R., & Watson, R. T. (2020). *Maritime Informatics*. Springer International Publishing.

Linke, M. (2012). Connecting Logistics Networks Globally Via the Un Single Window Concept. *Journal of Globalization Studies*, *3*(2), 139–154.

McKinsey. (2020). *From surviving to thriving: Business after coronavirus*. Author.

Mendes Constante, J., Langen, P., Vieira, G., Lunkes, R., & van der Lugt, L. (2018). The impact of management practices use on Brazilian Port Authorities' Performance. *Rivista Internazionale di Economia dei Trasporti*, *45*. Advance online publication. doi:10.19272/201806702005

Ministério da Infraestrutura. (2018). Porto sem Papel - PSP — Português (Brasil). Author.

Ministério da Infraestrutura. (2020). Sistema Portuário Nacional — Português (Brasil). *Governo Do Brasil*. Retrieved from https://www.gov.br/infraestrutura/pt-br/assuntos/transporte-aquaviario/sistema-portuario

Notteboom, T., Lugt, L. van der, Saase, N. van, Sel, S., & Neyens, K. (2020). The Role of Seaports in Green Supply Chain Management: Initiatives, Attitudes, and Perspectives in Rotterdam, Antwerp, North Sea Port, and Zeebrugge. *Sustainability, 12*(4), 1688. doi:10.3390u12041688

Notteboom, T. E., Parola, F., Satta, G., & Pallis, A. A. (2017). The relationship between port choice and terminal involvement of alliance members in container shipping. *Journal of Transport Geography*, *64*, 158–173. doi:10.1016/j.jtrangeo.2017.09.002

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

OECD. (2019). *OECD Public Governance Reviews The Innovation System of the Public Service of Brazil An Exploration of its Past, Present and Future Journey*. OECD Publishing.

Osobajo, O. A., Koliousis, I., & McLaughlin, H. (2021). Making sense of maritime supply chain: A relationship marketing approach. *Journal of Shipping and Trade*, *6*(1), 1. doi:10.118641072-020-00081-z

Parola, F., Satta, G., Notteboom, T., & Persico, L. (2020). Revisiting traffic forecasting by port authorities in the context of port planning and development. *Maritime Economics & Logistics*. Advance online publication. doi:10.105741278-020-00170-7

Portugal.EU. (2021). *Program for the Portuguese Presidency of the Council of the European Union*. Author.

Portuguese Government Portal. (2021). *Americas - Regional Affairs - Foreign Policy - Diplomatic Portal. Ministry of Foreign Affairs of Portugal*. Retrieved from https://www.portaldiplomatico.mne.gov.pt/en/foreign-policy/regional-affairs/americas

Projeto JUL- LSW contributes to the concept of the National Single Window. (2021). Retrieved from https://www.projeto-jul.pt/en/news/lsw-contributes-to-the-concept-of-national-single-window

Rodrigue, J.-P. (2010). *Maritime Transportation: Drivers for the Shipping and Port Industries*. Academic Press.

Rodrigues, K. R., Ferreira, C. G., Murta, A. L. S., & Murta, M. P. A. (2017). Sistema Portuário Brasileiro e o Uso da Tecnologia para uma Gestão Eficiente. *HOLOS*, (7), 110–126. doi:10.15628/holos.2017.6080

Srour, F., Oosterhout, M., Baalen, P. J., & Zuidwijk, R. (2008). *Port Community System Implementation: Lessons Learned from International Scan*. Academic Press.

Stopford, M. (2009). *Maritime Economics*. Routledge.

The Economist Intelligence Unit. (2020). *The Internet of Things: Applications for Business. Exploring the transformative potential of IoT*. Author.

UNCTAD. (2019). *Digitalization in Maritime Transport: Ensuring Opportunities for Development | UNCTAD*. United Nations Publication.

UNCTAD. (2020). Review of Maritime Transport 2020. *United Nations Publications*.

UNCTAD. (2021a). *Maritime profile: Brazil*. Retrieved from https://unctadstat.unctad.org/CountryProfile/MaritimeProfile/en-GB/076/index.html

UNCTAD. (2021b). Maritime profile: Portugal. *UNCTADstat*. Retrieved from https://unctadstat.unctad.org/CountryProfile/MaritimeProfile/en-GB/620/index.html

Vanelslander, T., & Sys, C. (2020). *Maritime Supply Chains*. Elsevier Science.

Vedan, A. B., & Thomas, E. M. B. (2019). A Cabotagem Brasileira no Contexto Mundial: Uma Análise Compreensiva Da Lei N° 9.432/1997 e de seus desdobramento. In E. M. O. Martins & P. H. R. de Oliveira (Eds.), *Direito marítimo, portuário e aduaneiro: temas contemporâneos* (pp. 13–31). Arraes Editores.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Venkatesh, V. G., Zhang, A., Deakins, E., Mani, V., & Shi, Y. (2020). Supply Chain Integration Barriers to Port-Centric Logistics—An Emerging Economy Perspective. *Transportation Journal*, *59*(3), 215–253. doi:10.5325/transportationj.59.3.0215

Vieira, J., & Fialho, G. (2020). Port Management Modernization and Integrated Operational Planning. *Revista Eletrônica de Estratégia & Negócios*, *13*, 196. doi:10.19177/reen.v13e0I2020196-224

Ward, R., & Bjørn-Andersen, N. (2021). The Origins of Maritime Informatics. In Maritime Informatics. Springer. doi:10.1007/978-3-030-50892-0_1

WCO. (2021). *Managing Transition to a Single Window, Volume 1, Part IX*. World Customs Organization.

World Shipping Council. (2021). *Industry Issues*. Retrieved from https://www.worldshipping.org/industry-issues

KEY TERMS AND DEFINITIONS

Bill of Lading: It is a document issued by the carrier that defines the contracting of the international transport operation and proves the receipt of the goods at the origin and the obligation to deliver them at the destination.

Digital Port: The evolution of traditional ports, which are converted into technological parks and innovation environments. Technology plays a significant role in them.

Digital Ship: A set of various trends in the digitalization of shipping companies.

Digital Supply Chain: The evolution of the supply chain, adapted to the use of technological resources and encouragement of innovative ideas from the manufacture of a product to its delivery to the final consumer.

IMO Number: Numeric field with seven digits corresponding to the code of the vessel according to identification from the International Maritime Organization (IMO).

Logistics Single Window (JUL – Janela Única Logística): This is a digital ecosystem that allows data sharing between port authorities and other national authorities and organizations at the level of the European Commission.

Paperless Port (PsP – Porto sem Papel): The initiative promoted by the Federal Government when the system automated the provision of information necessary for the stay of vessels in Brazilian public ports.

Port Authority: Or Port Administration. Is the entity with jurisdictional power in a specific location to operate ports.

Public Port: Public good built and equipped to meet the needs of navigation, passenger movement, or the movement and storage of goods, and whose traffic and port operations are under the jurisdiction of the port authority.

TEU – Twenty-Foot Equivalent Unit: Refers to the Equivalent Transport Unit. This transport unit has a 20-foot intermodal container size. A 40-foot container corresponds to 2 TEU.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 2
A Software Engineering Perspective on Building Production-Ready Machine Learning Systems

Petra Heck
Fontys University of Applied Sciences, The Netherlands

Gerard Schouten
Fontys University of Applied Sciences, The Netherlands

Luís Cruz
Delft University of Technology, The Netherlands

ABSTRACT

This chapter discusses how to build production-ready machine learning systems. There are several challenges involved in accomplishing this, each with its specific solutions regarding practices and tool support. The chapter presents those solutions and introduces MLOps (machine learning operations, also called machine learning engineering) as an overarching and integrated approach in which data engineers, data scientists, software engineers, and operations engineers integrate their activities to implement validated machine learning applications managed from initial idea to daily operation in a production environment. This approach combines agile software engineering processes with the machine learning-specific workflow. Following the principles of MLOps is paramount in building high-quality production-ready machine learning systems. The current state of MLOps is discussed in terms of best practices and tool support. The chapter ends by describing future developments that are bound to improve and extend the tool support for implementing an MLOps approach.

DOI: 10.4018/978-1-7998-6985-6.ch002

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

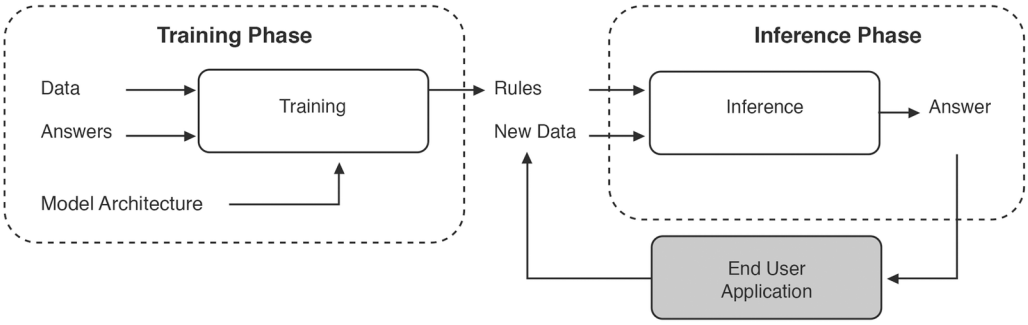
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The application of data science and artificial intelligence (AI) in business and industry is not a niche anymore. AI is used in traditional areas like machine vision, speech recognition, and translation, and in a wide variety of novel areas like detecting fraud in transaction data, decoding and identifying hand-written text, medical diagnosis, or even tracking wildlife. Machine learning (ML) is seen as a subset of AI. The term *machine learning* denotes a set of algorithms that learn from data. Machine learning (ML) also includes Deep Learning (DL), which denotes a set of algorithms that use multi-layered neural networks to learn from data. Currently most AI implementations are ML implementations. In order to build solutions that can be delivered to customers, ML should be connected to software. The software solution ensures that what the ML model learns from the data is transformed into meaningful predictions or decisions for the end-user. The need for such software solutions is growing fast as the number of AI applications increases. This chapter discusses software solutions that contain an ML component and focuses on the engineering approach required to construct them. The remainder of this chapter refers to this type of software solutions as "ML systems", to indicate that they consist of both an ML model and a software solution or software system.

Figure 1 illustrates the relevant concepts of an ML system. This chapter focuses on supervised ML, where a model needs to be trained on historical data, labeled with answers. After a model has been trained successfully, it must be deployed somewhere such that a software solution can feed it new data and retrieve answers for this new data. This is called inference in ML terminology.

Figure 1. Incorporating a (supervised) ML model into a software application (adapted from Cai et al., 2020)



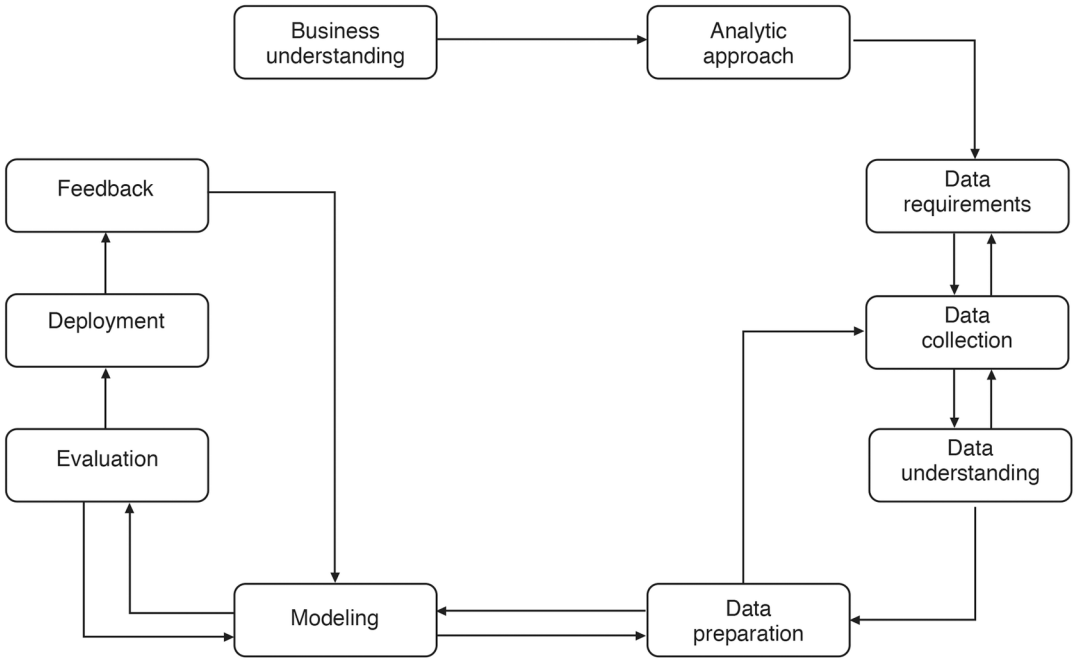
As the application of ML in business and industry matures, more and more organizations reach the point where they need to run an ML software solution in their production environments. Yet, the data scientists designing the models have not been trained in the software engineering skills required to put their models into production. On the other hand, the addition of ML components introduces some new challenges for software developers building the solution.

This chapter discusses those challenges and provides the solutions (in terms of good engineering practices and supporting tools) available to date. The chapter starts by providing some background on

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the development process and maturity levels of ML systems. After a discussion of individual challenges and solutions, the chapter introduces MLOps as an overarching approach to build production-ready ML systems. The chapter ends by discussing related work and future directions.

Figure 2. Steps in a data science project according to IBM Data Science Methodology (adapted from Rollins, 2015)



BUILDING PRODUCTION-READY ML SYSTEMS

This section elaborates on the development process of ML systems and defines ML engineering. This section also introduces different levels of maturity for ML systems.

Stages in ML Engineering

The diagram depicted in Figure 1 shows that building an ML model consists of a training phase and an inference phase. The training phase is what data scientists mainly focus on. However, from an engineering perspective, much more needs to be done to create a reliable ML system. A practical and widely adopted approach for translating the engineering perspective into a number of logical and manageable project steps is the IBM Data Science Methodology (Rollins, 2015), based on CRISP-DM (Chapman et al., 2000).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The IBM Data Science Methodology (Rollins, 2015) proposes a schematic overview of a complete data science project. As depicted in Figure 2, it consists of the following steps:

- **Business understanding:** define the problem, project objectives and solution requirements from a business perspective.
- **Analytic approach:** express the problem in the context of statistical and ML techniques so that the data scientist can identify techniques suitable for achieving the desired outcome.
- **Data requirements:** based on domain knowledge, determine data content, formats, and representations required for the analytic methods.
- **Data collection:** identify and gather data resources - structured, unstructured, and semi-structured - that are relevant to the problem domain.
- **Data understanding:** use descriptive statistics and visualization techniques to understand data content, assess data quality, and discover initial insights into the data.
- **Data preparation:** construct the dataset that will be used in the modeling stage. This includes data cleaning, combining data from multiple sources, and transforming data into more useful variables. Moreover, feature engineering and text analytics may be used to derive new structured variables.
- **Modeling:** use a training set - historical data in which the outcome of interest is known - to develop predictive or descriptive models using the analytic approach already described. The modeling process is highly iterative.
- **Evaluation:** evaluate the model's quality and check whether it addresses the business problem fully and appropriately. Doing so requires the computing of various diagnostic measures - as well as other outputs, such as tables and graphs - using a test (data) set for a predictive model.
- **Deployment:** after a satisfactory model has been developed that has been approved by the business sponsors, deploy it into the production environment or a comparable test environment.
- **Feedback:** get feedback on the model's performance and observe how it affects its deployment environment. Analyze this feedback to refine the model, increase its accuracy, and thus its usefulness.

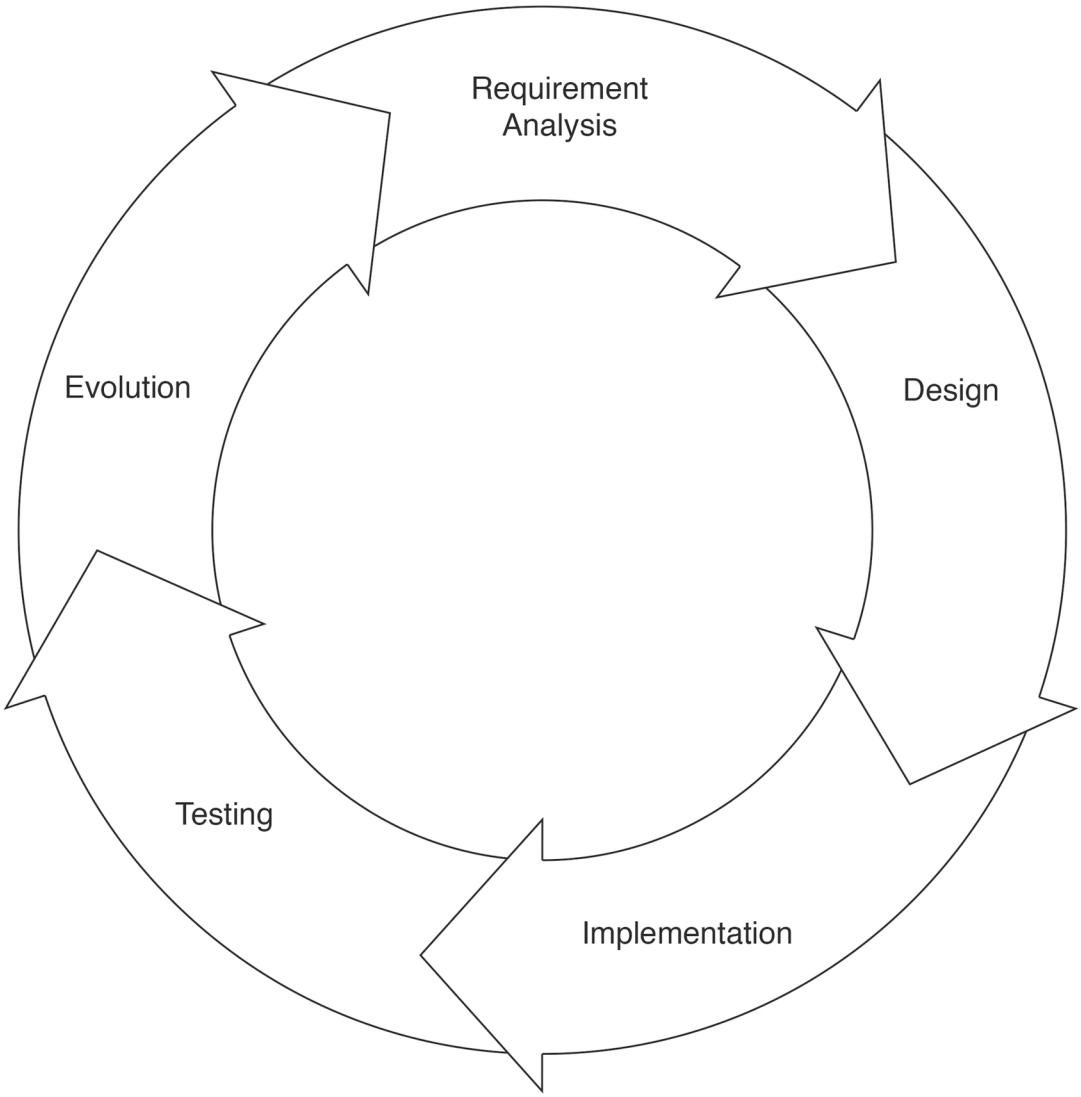
The IBM process shows that modeling can only start after a good deal of data engineering. More importantly, the data science cycle should always start with a good understanding of the ML model's business goal(s). When a good enough model is found, the model needs to be deployed (inference phase) and continuously monitored.

The IBM view on ML projects works well for the ML modelling side of the project. Yet, it does not provide much guidance for the development of the accompanying software solution, let alone the deployment of the complete ML system in a production environment (Rollins, 2015). For software development, many different lifecycle models exist, but all include – in slightly different forms – five major activities that are fundamental to software engineering (see Figure 3):

- **Requirement analysis:** define the problem, project objectives and software requirements from a business perspective (i.e., business understanding). These requirements can be functional (what does the software need to do for the user?) or non-functional (how does the software need to do this for the user, e.g., concerning performance, usability, and security).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. A standard software development lifecycle (adapted from Cliffydcw, 2020)



- **Design:** translate the software requirements into a componentized design for the software structure, also called software architecture. This involves decomposition to break the software into smaller modules with a well-defined responsibility and abstraction to avoid implementing the same structure twice.
- **Implementation:** build the software as specified in the design.
- **Testing:** evaluate the software quality and check whether it addresses the software requirements fully and appropriately.
- **Evolution:** after a satisfactory software product has been developed and approved by the business sponsors, deploy it into the production environment or a comparable test environment. Get

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

feedback on the performance and observe how it behaves and affects its deployment environment. Analyze this feedback to refine the software and increase its usefulness.

Although the software development lifecycle in Figure 3 has similarities with the ML process in Figure 2, there are also two important differences:

1. An ML project starts with data requirements, whereas a software development project starts with software requirements.
2. Software development has an explicit design phase where structured methods can be used to derive software designs from requirements. In contrast, the design of ML models has a more experimental character and is intertwined with the data engineering and modeling steps.

Software engineers and data scientists need to create a common business understanding because, in the end, their products (model and code) will be deployed together as one ML system for the business. They should start preparing this deployment and how to get feedback after deployment, together, right from the start of the project.

Another term used for this combination of ML and software engineering is "ML Engineering". Burkov (2020) defines ML engineering as "the use of scientific principles, tools, and techniques of machine learning and traditional software engineering to design and build complex computing systems". This chapter further details the practice of ML engineering (i.e., building production-ready ML systems) by discussing challenges and solutions.

Maturity of ML Systems

In practice, many ML projects do not start from an engineering perspective. They have a more experimental character. Best practices in ML engineering depend on the maturity of the ML system. Some examples of ML maturity models are:

- Lwakatare et al. (2019) identified five evolution stages for the use of ML components in software systems: *experimentation & prototyping, non-critical deployment, critical deployment, cascading deployment, autonomous ML components*. For each stage, they identify different types of challenges.
- Akkiraju et al. (2020) define a re-interpretation of the software Capability Maturity Model (CMM) for the ML model development process, which has maturity levels *initial, repeatable, defined, managed,* and *optimizing*. For each maturity level, ML engineering capabilities are described and worked out.
- The Institute for Ethical AI & Machine Learning (2020) published a maturity model with a list of assessment criteria for AI solutions. The maturity model does not have explicit maturity levels.
- Algorithmia (2020) presents a roadmap to ML maturity. They define practices on three maturity levels (*low, medium, high*).

All these maturity models have in common that they acknowledge the fact that 1) ML engineering is an organization-specific process; 2) the path towards this process consists of implementing a set of best practices targeted at a specific maturity level for specific lifecycle activities. The remainder of this

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

chapter should also be seen in this light. Some best practices discussed might not fit the targeted maturity level of the project or organization. It is key, first to determine the targeted maturity level, and then choose the best practices that help most in attaining that maturity level.

ADDRESSING CHALLENGES FOR BUILDING ML SYSTEMS

Including ML models in a software system leads to several challenges (Heck, 2019). This section discusses five of those challenges, which play an important role in understanding what is needed for production-ready ML systems. For each challenge, a number of best practices and supporting tools are discussed that address the challenge.

[CH1] Elicitation of Data and Model Requirements

As shown in Figure 2, the ML lifecycle starts with the "Analytic Approach" and "Data Requirements/Collection/Understanding/Preparation". The analytic approach determines which type of ML model would be suited for the problem at hand, and the data steps aim to collect and prepare the data needed to train the model. It is important to understand that the trained model's quality depends on the data examples used for the training. In a supervised training setting, the model cannot learn patterns that are not present in the training data. On the other hand, the model might learn overly detailed patterns only present in the training data and do not generalize to new data (this is called overfitting in ML theory). Whereas at the start of software projects a lot of time is spent on understanding the user and the user requirements (how will they use the software, which tasks need to be supported, which user interface is suitable), in ML projects, considerable time is spent on understanding the data and the data requirements.

Next to the data requirements, an ML system also has model requirements. Vogelsang and Borg (2019) highlight that for ML projects, some functional software requirements need to be restated in terms of ML performance measures like accuracy. It might be difficult to find a link between those technical performance measures and the effect for the end user. They also describe some new non-functional requirements such as explainability, freedom from discrimination, or specific legal requirements that might come into play. This last point is reiterated by J. Zhang et al. (2020), who compiled a list of eight quality attributes to test for in ML systems: correctness, overfitting degree, robustness, security, data privacy, efficiency, fairness, and interpretability. These issues are, in practice, often neglected. For instance, Abbasi-Sureshjani et al. (2020) found in a survey that in most medical ML models for diagnosis, demographics are rarely explicitly taken into account, while in clinical practice factors such as gender and age should obviously be considered.

For an ML system, data, model, and software requirements should be elicitated. Both software engineering (*domain modeling* and *requirements engineering*) and data science (*exploratory data analysis, data quality check, data visualization*) provide approaches for that elicitation. Combining these five methods leads to high-quality requirements for all aspects of the ML system. High-quality requirements are the basis for building a high-quality (production-ready) ML system. The remainder of this section discusses each of these five solutions in more detail.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Domain Modeling

Most software solutions contain some form of a database with the data needed to support the end-user's tasks. To understand the key concepts and the relations between them, the software engineer uses techniques like Entity-Relationship (ER) modeling (Thalheim, 2013). Such an ER model can then be discussed with the domain expert. As an extension of this, Kühn et al. (2018) describe what they call the Analytics Canvas. This is a semi-formal specification technique for describing analytics use cases, the necessary data infrastructure, and data requirements for each domain. It uses a graphical notation to support communication between the various stakeholders of different domains in an analytics project.

Even if not using such explicit modeling techniques, it is paramount that in the data requirements phase of an ML project, a domain expert is available who can answer questions about the data generation process and can confirm intuitions, assumptions, and findings. Provost and Fawcett (2013) promote what they call "Data-Analytic Thinking". Their book helps business stakeholders (e.g., domain experts) to view business problems from a data perspective and understand principles of extracting useful knowledge from data. Business stakeholders faced with a business problem should be able to assess whether and how data can improve performance. Furthermore, if business stakeholders understand some fundamental data science concepts, this facilitates communication with data scientists.

Requirements Engineering

Requirements engineering has long been considered a separate discipline within the software engineering discipline (Bourque & Fairley, 2014). Sommerville and Sawyer (1997) present a systematic approach to requirements engineering applicable to any application type. They claim that use of a systematic approach based on engineering principles leads to better requirements than an informal approach. This is, of course, also valid for ML systems.

Exploratory Data Analysis

Understanding data requires an exploratory process where the ML engineer "goes through the data," finds something interesting, and goes back to the domain expert to confirm the findings or ask questions. This process should repeat until the ML engineer is confident enough that the data has no more secrets, and the quality of the data (correctness and validity) is sufficient to train a model on. This process is called Exploratory Data Analysis. Alspaugh et al. (2018) concluded from their survey that practitioners in the field really do "ask questions of their data" in an exploratory fashion. Their study also shows a pervasive use of programming languages, scripting tools, and visualization for exploration. Kery et al. (2018) found that a literate programming tool like Jupyter notebooks helps data scientists share exploration results with domain experts. Such a tool merges data manipulation code, data visualizations, and written explanations to let the domain expert follow and validate the data story.

Data Quality Check

Before using a dataset for further analysis, it is important to detect incomplete, inaccurate or irrelevant parts of the data. Burkov (2020) describes best practices to avoid common problems with data.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

For data testing, it is mainly up to the engineer to come up with good test cases. Preferably these test cases are automated such that the data quality check is easily repeated when new data comes in. Heck (2020) lists some additional tools and techniques that can be used for testing data.

Data Visualization

Graphics provide an excellent approach for exploring data and presenting results (Chen et al., 2007). The Handbook of Data Visualization (Chen et al., 2007) offers principles and methodology to create those graphics. Alspaugh et al. (2018) found that data scientists use all sorts of different tools for data visualization: code in programming tools like Jupyter notebooks, database-related tools, spreadsheet tools like Excel, interactive visualization toolkits like Tableau.

Keim et al. (2010) describe the evolution of visual data exploration into the discipline of *visual analytics*. According to them, "visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets." Their book provides an overview of the discipline and provides many pointers to interactive visualization techniques and tools.

An ML engineer should become skilled in data visualization. It is a useful tool in data and model exploration (Alspaugh et al., 2018), and it is a mandatory tool to communicate with business stakeholders (Chen et al., 2007).

[CH2] Modularizing the Application

As shown in Figure 3, software design plays an important role in developing software systems. Software design (Bourque & Fairley, 2014) principles like abstraction, decomposition, and encapsulation are used to organize the software system into components with high cohesion and low coupling. The components are specified at a level of detail that enables their construction. Splitting larger systems into components lowers the complexity of the development. Components can even be implemented and tested by separate teams.

Software designs are white-box (i.e., they specify the internal structure and behavior), whereas ML models, especially neural networks, are mainly black-box. Specifying an ML model is done by describing input and output data. Based on that, the selected ML algorithm learns its internal structure (i.e., the function that maps input to output). This means that an ML model cannot be divided into smaller components: it functions as a whole. Moreover, according to Amershi et al. (2019), maintaining strict module boundaries between machine learned models is difficult.

Software engineering provides several approaches to design high-quality applications, including *componentization, design patterns, pipelines, and containerization*. The remainder of this section describes how each of these four approaches applies to ML system design. After defining high-quality requirements, a high-quality design is the next step toward building high-quality (production-ready) ML systems.

Componentizing ML Systems

An important trend to improve ML systems' maintainability is to move away from monolithic end-to-end ML systems. As pointed out by Marcus and Davis (2019), especially deep learning systems are extremely hard to debug because of this end-to-end approach. Instead, they should be split-up into smaller logi-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

cal self-contained components. For instance, a more recent model architecture for autonomous driving promotes a separate ML component for perception (object detection and classification), followed by another ML component for near real-time decision making, and a third one for mapping out longer-term predictions. Although individual ML models cannot be divided in smaller components, the ML system as a whole should be structured to solve a large complex problem piece-by-piece.

ML Design Patterns and Standards

Washizaki et al. (2020) discuss architecture patterns, design patterns, and anti-patterns for ML application systems. By analyzing relationships between patterns, they identify "Separation of Concerns and Modularization of ML Components" as one of two fundamental patterns. The anti-patterns on their list are mostly taken from Sculley et al. (2014), which is a seminal paper in the area of engineering ML applications. McAteer (2020) discusses this paper by Sculley et al. (2014) in light of the newest technologies to address and avoid the anti-patterns.

Pinhasi (2020) introduces architecture patterns for decomposing inference tasks. An inference pipeline (i.e., pipeline to get predictions on new data from a trained model) consists of a business interface, a module to translate business objects to model data, a module to predict, and a module to translate predictions to business outcomes. Services are a relatively new way in software engineering to componentize software systems (Kang et al., 2016). Pinhasi describes patterns of how to deploy the inference pipeline into production as one or several services.

The ISO/IEC JTC 1 (Joint Technical Committee 1) develops, maintains, and promotes standards in the fields of Information Technology (IT) and Information and Communications Technology (ICT). Subcommittee 42 does this for the area of Artificial Intelligence. They have published a reference architecture for Big Data and are working on what they call a "reference architecture of knowledge engineering". In the future, they might also come up with ways to structure the design of ML systems.

Pipelines

A concept widely used in ML applications is called *pipeline*. A pipeline automates an independent workflow in the ML project. The tasks in the workflow are steps in the pipeline. By creating separate pipelines for, e.g., data preparation, configuration, training, and deployment, the whole ML lifecycle can be modularized. Steps within a pipeline are even smaller "modules" of the overall process. Independent steps allow multiple data scientists to work on the same pipeline at the same time. Steps can be tested separately and even re-used for other projects. The importance of pipelines is also recognized by Amershi et al. (2019): "It is important to develop a 'rock solid' data pipeline capable of continuous loading and massaging data, enabling engineers to try out many permutations of AI algorithms with different hyper-parameter without hassle." For example, the Microsoft Azure Cloud environment offers pipelines as a standard concept for developers (Microsoft, 2020a).

Containerization

Benton (2020) describes the concept of intelligent applications (i.e., ML systems) and proposes to structure them on a containerization platform like Kubernetes (kubernetes.io). Containers have emerged in cloud environments to provide a light-weight virtualization solution for application packaging (Kang et

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

al., 2016). Benton presents an idealized intelligent application architecture where the system is divided into multiple containers that can be developed, tested, deployed, and updated in isolation. Kubernetes is then used for the orchestration of these isolated containers. Benton also mentions Kubeflow (kubeflow.org) that provides additional tooling to make ML scalable and repeatable on Kubernetes.

[CH3] Design through Experimentation

Although Figure 2 gives the impression that the ML process is nicely structured into sequential steps, in reality, it is highly experimental and iterative by nature. Ameisen (2018) describes the ML engineering loop with the following iterative steps: *analyze, select an approach, implement,*

measure. He observes that it is a challenge for ML teams to maintain a level of forward progress due to the development process's uncertainty. Sommerville (2019) sees similar challenges in what he calls "exploratory development": *formulate ideas, build a prototype,* and *assess ideas*. He contrasts this to the specify-build-validate approach that he believes is deeply embedded into many organizations. Burkov (2020) similarly specifies model building as an iterative process where an analyst builds a model, observes its behavior, and adjusts the model based on the observations. His ML project lifecycle contains many feedback loops.

ML systems need high-quality development processes to deal with this uncertainty. Software engineering has introduced development approaches and tools to deal with the uncertainty that also transfer to a data science process (*agile data science*, *development tools for experiments* and *AutoML*). Each of these three solutions is detailed in the remainder of the section. An agile approach and automation lead to high-quality development processes that are needed to build high-quality (production-ready) ML systems.

Agile Data Science

In software engineering, the way to deal with uncertainty has been to switch to so-called agile development methods. Agile methods are characterized by short, iterative development cycles, self-organizing teams, simpler designs, code refactoring, test-driven development, frequent customer involvement, and an emphasis on creating a demonstrable working product with each development cycle. Jurney (2017) translates these agile principles to the ML environment. His book also includes a technical implementation with tools for each of the steps. Yan (2019) also concludes that agile and data science are a good match. His post includes some frameworks for effectively applying agile to data science. The Microsoft Team Data Science Process (Microsoft, 2020b) is a framework developed by Microsoft that provides a structured sequence of activities to develop cloud-based predictive analytics solutions efficiently. The Team Data Science Process includes specific workflows and templates for agile development.

Ameisen (2018) discusses each step in the engineering loop (Analyze, Select Approach, Implement, Measure). In line with agile methods, he recommends getting started with a minimal implementation and improve from thereon. Next to improving the model's performance in each iteration, the quality and speed of the iteration itself can also be improved.

Development Tools for Experiments

For the ML engineer to set up and track experiments, tool support is essential. A lot of such tools are being built right now, most of them as open source:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Literate programming tools, e.g., Polynote (polynote.org), that includes support for data visualizations without programming.
- Libraries, e.g., PyCaret (pycaret.org), that allows training and deployment of supervised and unsupervised ML models in a low-code environment.
- Platforms for experiment tracking and automation of the ML lifecycle, e.g., MLFlow (mlflow.org) and Comet (comet.ml).

Automated ML (AutoML)

Another option to conduct experiments faster is to automate tasks. In recent years a number of platforms have been launched that perform such automated ML (AutoML or AutoAI). Examples are Google AutoML (cloud.google.com/automl/docs), H2O Driverless AI (h2o.ai), and RapidMiner Auto Model (rapidminer.com). Such platforms allow business users with limited knowledge of ML to create ML models by uploading training datasets. These platforms use automated model selecting, automated feature engineering, and automated hyperparameter tuning to generate hundreds of possible models. The end user has a graphical UI (no programming) to interact with the platform and select the best scoring model. Wang et al. (2019) conclude that AutoML is a valuable addition to the data science landscape that will augment data scientists' work, not replace it.

[CH4] Data and Model Management

Version control, requirements management, and configuration management are standard practices in software engineering projects (Bourque & Fairley, 2014). Those practices aim to track and control changes in the software: code, design, or requirements. For software projects, the main artefact that needs to be tracked is the source code. Platforms like Git (git-scm.com) provide a distributed version-control solution for tracking changes in source code during software development. ML systems introduce two other important artefacts for which changes need to be tracked: the dataset and the ML model. Traditional version-control systems work file-based, whereas the ML engineer needs a more fine-grained notion of version control (Kery et al., 2018). For tracking experiments, it is needed to easily see, for example, which columns or rows were added in the dataset or which parameter setting was altered for the ML model. Furthermore, the ML engineer needs a system in which the relationship between data, model, and code can be versioned: which dataset was used with which data transformation code with which ML algorithm with which parameter settings with which glue code to produce which results?

Another challenge regarding the management of data and models is knowing when and how to reuse existing models for a new project. It is currently not common practice to document datasets or models such that others can understand their requirements and constraints (Mitchell et al., 2019). Understanding if the old data is still valid for the new project or if it is worth investing time in retraining an existing model is complicated. Often, the easiest way to draw conclusions on reuse is to try out if it works, but this might take a considerable amount of time. The other way around, due to the ubiquitous availability of open-source datasets and models online, it might take a long time to find something fit for reuse in the new project. Most organizations currently do not have searchable model repositories.

High-quality ML systems need to be traceable in all aspects: data, model and code. Software engineering provides approaches for traceability and reuse that also translate to ML systems: *version control*, *documentation, APIs, and libraries*. The remainder of this section discusses the application of each of

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

these three approaches to ML systems in more detail. Having traceable and properly documented artefacts is a prerequisite for putting an ML system into production, even more so for updating the ML system once it is in the production environment.

ML Version Control

As traditional version control tools are not suited for version control of data and models, new tools arise. A tool like DVC (dvc.org) is built on top of Git (git-scm.org) to extend version control to ML models and datasets. It uses automatic metric-tracking to manage experiments. DVC is more data-oriented and often used together with ML lifecycle tools like MLFlow (mlflow.org), providing a central model store to collaboratively manage the full lifecycle of an ML model, including model versioning, stage transitions, and annotations. Cloud ML environments like the before mentioned Microsoft Azure (Microsoft, 2020a) extend their pipeline concept with data stores (Azure ML Datasets) and model stores (Azure ML Model Registry) to support tracking of all artefacts, including pipelines.

IT-advisory firm Gartner (Krensky et al., 2020) notes the rise of what they call DSML (Data Science and ML) platforms like Dataiku (dataiku.com) and Databricks (databricks.com) that include data and model version control as part of an integrated whole, avoiding the somewhat difficult setup of separate tools like DVC, MLFlow and Azure ML.

Data and Model Documentation

An obvious way of making it easier for ML engineers to assess if a model or dataset is fit for reuse is to document them better. Zinkevich (2017) recommends documenting each relevant column of the dataset with an owner, a detailed description of the column, where it came from, and how it is expected to help. Mitchell et al. (2019) propose the use of what they call "Model Cards." Model cards are short documents accompanying trained ML models that provide benchmarked evaluation in a variety of conditions relevant to the intended application domains. Model cards also disclose the context in which models are intended to be used, details of the performance evaluation procedures, and other relevant information. Kriens and Verbelen (2019) propose to specify the ML model's capabilities using the OSGi specification for Java-based applications (osgi.org). However, none of these documentation options are standard practice yet.

Model APIs and Libraries

An easy way to reuse ML models built by others is through the use of a model application programming interface (API). An API is common in software engineering; it abstracts the underlying implementation and only exposes objects or actions the developer needs. In the case of a model API, the ML engineer sends data to the API, and the API returns the predicted labels. An example of this is Microsoft Cognitive Services (Raman & Hoder, 2020) that offers APIs in five categories: vision, speech, language, decision, and web search. These APIs are offered for generic tasks, like detecting objects in photographs, with models that have been trained with huge datasets. For specific tasks, like detecting lice on a flower stem, these APIs will not perform well. However, more and more APIs and other packages appear for such specific tasks. Examples are PoseNet (Oved, 2018), which provides a model for real-time human pose estimation, and OpenCV (opencv.org), a library of programming functions mainly aimed at real-time computer vision.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

[CH5] Testing

Software testing is a separate lifecycle phase with ample best practices and tool support (Bourque & Fairley, 2014). For ML systems, both data, model, and code need to be validated. J. Zhang et al. (2020) describe the testing workflow (how to test), testing properties (why to test), and testing components (where and what to test) for ML systems. They also describe how ML testing differs from testing traditional rule-based software. The properties to test for that they identify are: correctness, overfitting degree, fairness, interpretability, robustness, security, data privacy, and efficiency. After this verification, the complete ML system should be validated with the end user: are the ML model results as embedded in the system meaningful and useful for the end user? In software engineering, this validation is called "user acceptance testing" (Bourque & Fairley, 2014). User acceptance testing for ML systems is more difficult because the system is designed to provide an answer to a question for which no previous answer exists (J. Zhang et al., 2020).

The remainder of this section discusses several approaches from software testing and how they translate to ML systems: *testing ML systems, online testing, test set adequacy, test-driven ML,* and *testing for interpretability and fairness)*. Having a high-quality test approach is paramount to determine if the ML system is, in fact, production-ready, and for monitoring the quality of the ML system after deployment in the production environment.

Testing ML Systems

The A4Q AI and Software Testing Foundation certification includes a syllabus on testing ML applications and the use of ML in testing (Alliance for Qualification, 2019). The certification is related to the ISTQB Certified Tester (istqb.org), which is the global standard in Software Testing.

J. Zhang et al. (2020) provide a survey of 128 papers on testing ML. Heck (2020) builds on this survey to supply a list of practical engineering tools and best practices for testing ML systems. One of the best practices mentioned there, which is often overlooked, is the use of static code analysis to ensure high-quality code. This is standard practice in software engineering with dedicated static analysis tools (Bourque & Fairley, 2014), but currently, no static analysis tools have been developed specifically for ML code. At least one should define coding standards or coding practices for ML code (Tan, 2019) and enforce a manual check against those coding standards.

Online Testing and Monitoring

An ML testing workflow should also contain online testing activities to complement the shortage of offline testing (J. Zhang et al., 2020). Online testing collects metrics of model performance in the production environment. A common scenario is to monitor user behavior in response to different model versions (Burkov, 2020). Popular techniques for this are A/B testing and multi-armed bandit (Burkov, 2020).

A requirement for online testing is that continuous monitoring of the production model(s) is in place. Burkov (2020) describes what needs to be monitored for ML systems and how to do it. Hulten (2019) puts it more abstract by stating it is important to "know how to craft experiences that collect the data needed to evaluate and grow intelligence." His book gives some non-technical guidance with chapters on validation, evaluation, and orchestration of intelligent systems.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Test Set Adequacy

The goal of test adequacy evaluation is to determine whether the existing tests have a good fault-revealing ability, and provide an objective confidence measurement on the testing activities (J. Zhang et al., 2020). In traditional software testing, code coverage is used to determine the adequacy of a test set. Code coverage indicates the percentage of lines of code that are covered while executing a test set. Such a metric does not entirely fit ML projects. Thus, some adaptions of code coverage have been proposed. Burkov (2020) describes neuron coverage as a variant for deep learning models. He also describes a variant of mutation testing where the training data is mutated instead of the software code. J. Zhang et al. (2020) describe a few more coverage notions, all for deep neural networks. They note the limitations of these notions in practice. Breck et al. (2016) present a set of rules and offer a scoring system to measure how ready for production a given ML system is. Their rules could also be used to check the completeness of the test set.

Test-Driven ML

Test-driven development (TDD) means that tests are written before the code. Adopting TDD moves the focus on validation to the beginning of the project, which leads to higher-quality systems. This technique aims to reduce the fear of making changes in software projects. Hence, it promotes refactoring without having a fear of breaking things. Since ML projects tend to be very experimental, TDD is quite handy for ML projects as well. Bozonier (2015) and Kirk (2017) provide detailed examples in their books that help practitioners adopt test-driven ML. Despite requiring major behavior changes in the development practices, test-driven ML is expected to gain popularity. It encourages having a complete and up-to-date set of tests at different ML system levels: code, algorithm, and project.

Testing for Interpretability and Fairness

As said before, J. Zhang et al. (2020) list eight different test properties for ML systems. Out of those eight properties, interpretability and fairness have received considerable attention in terms of solutions and best practices, of which this section discusses a few examples.

ML interpretability refers to the degree to which an observer can understand the cause of a decision made by an ML system (J. Zhang et al., 2020). A side-effect is that interpretable models are easier to test for correctness because they provide more means to find root causes of incorrect answers (Ribeiro et al., 2016; Zinkevich, 2017). Another term used for interpretable models is *Explainable Artificial Intelligence (XAI)*. Adadi and Berrada (2018) provide an organized overview of the existing XAI approaches, based on a literature analysis of 381 papers. Heck (2020), in addition, provides links to a book on "Interpretable ML" with examples in R (Molnar, 2020) and to Python libraries that implement some of the well-known interpretation methods. The Institute for Ethical AI & ML published the explainability toolkit XAI, available for Python on GitHub (github.com/EthicalML/XAI). Microsoft (Nori et al., 2019) published the open-source toolkit InterpretML (interpret.ml), a Python package to help understand models and enable responsible ML.

A fairness bug occurs when differences in model performance are measured for distinct groups or individuals. J. Zhang et al. (2020) summarize the work on fairness for ML, including a few test generation techniques for fairness testing. IBM published the AI Fairness 360 Open Source Toolkit (Bellamy

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

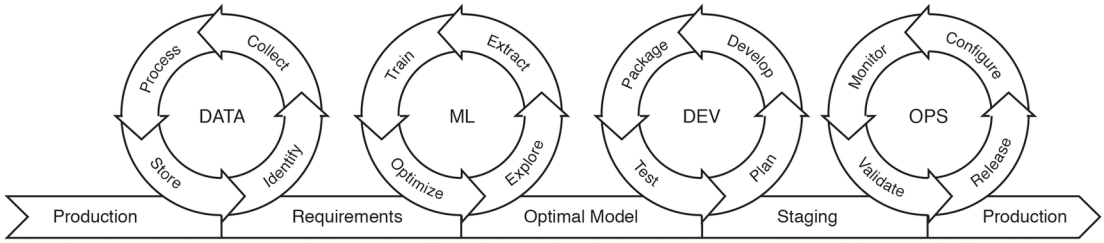
et al., 2019). The AI Fairness 360 Python package includes a comprehensive set of metrics for datasets and models to test for biases, explanations for these metrics, and algorithms to mitigate bias in datasets and models. Microsoft (Bird et al., 2020) published the FairLearn toolkit (fairlearn.github.io), a Python package with capabilities to assess the ML models' fairness and to mitigate unfairness.

Interpretability and fairness are both intricately connected to "Ethical AI" (Floridi & Cowls, 2019). Spiekermann (2015) introduces the Ethical System Development Lifecycle (E-SDLC) and provides a tool-supported methodology for value-based IT Design. A more accessible alternative is the Technology Impact Cycle Tool (tict.io), which asks questions to help estimate the impact of (new) technology. It includes questions on fairness and interpretability.

MLOPS: A UNIFIED APPROACH TO BUILDING ML SYSTEMS

DevOps (Ebert et al., 2016) emerged in the software community as a way of combining all aspects of software development and software maintenance into a set of best practices. Recently, the extension to ML systems has been coined MLOps (Visengeriyeva et al., n.d.). As illustrated in Figure 4, MLOps proposes an integrated approach that adds data and ML artefacts to the traditional DevOps approach.

Figure 4. MLOps blueprint, (adapted from Farah, 2020)



Such an integrated lifecycle approach entails several challenges, as discussed in the previous section. This section discusses the solutions brought upon the community by the MLOps movement.

The MLOps Process

According to Visengeriyeva et al. (n.d.), MLOps should provide the following capabilities on top of DevOps:

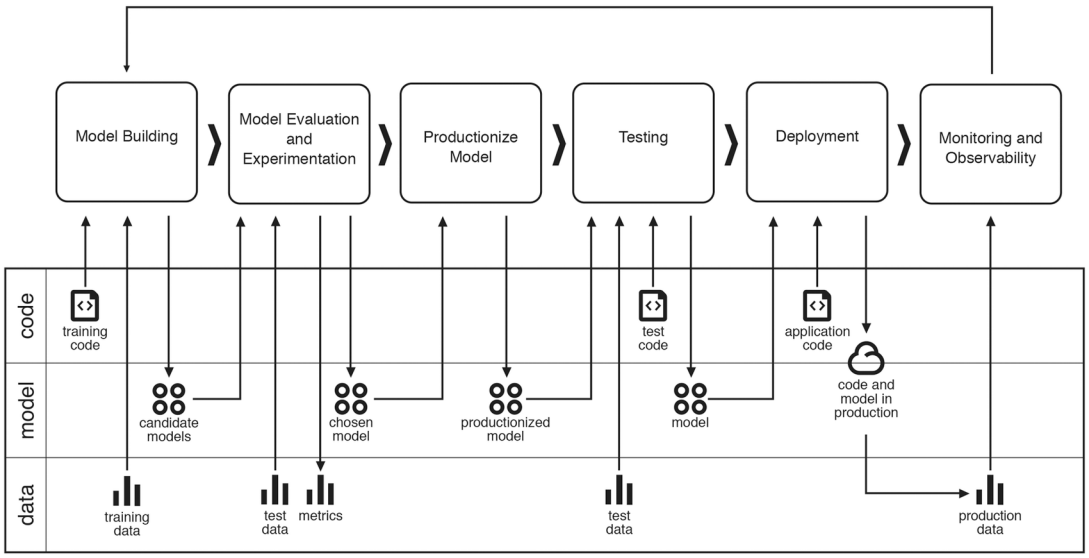
- Handle ML models and datasets as first-class citizens within Continuous Integration and Continuous Delivery (CI/CD) systems.
- Unify the release cycle for ML and software application release.
- Leverage automated testing of ML artefacts.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Apply agile principles to ML projects.
- Reduce technical debt across ML models.
- Be a language-, framework-, platform-, and infrastructure-agnostic practice.

CI/CD is a practice from agile software development that automates testing, integration, and deployment of software in such way that a new version of the software can be available in the production environment more quickly and more frequent. As recognized by Visengeriyeva et al. (n.d.), CI/CD is also essential for ML applications. Accordingly, Sato et al. (2019) have proposed Continuous Delivery for ML (CD4ML) – "a software engineering approach in which a cross-functional team produces ML applications based on code, data, and models in small and safe increments that can be reproduced and reliably released at any time, in short adaptation cycles". The CD4ML is illustrated in Figure 5, with all the different artefacts (data, model, and code) in an ML project with respect to the lifecycle stages in which they are involved.

Figure 5. CD4ML end-to-end process, (adapted from Sato et al., 2019)



Google Cloud (2020) extends the CI/CD concept by proposing Continuous Training (CT): automatically retraining and serving ML models in the production environment whenever new data comes in, without or with little human intervention. The same work defines three maturity levels for MLOps approaches, depending on the level of automation of ML projects:

1. **MLOps Level 0:** Integration and deployment are done manually.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. **MLOps Level 1:** Both the model and the pipeline used to train it are deployed in the target environment. This allows for CT: the training pipeline runs in production to generate new models as soon as new data comes in.
3. **MLOps Level 2:** CI/CD is enabled. Any change in the ML pipeline (e.g., feature engineering, model architecture, hyperparameters) will automatically build, test, and deploy the new pipeline components to the target environment.

Depending on the type of project, different levels of maturity should be considered. For example, level 0 is ideal for small exploratory projects that only have a few versions through their lifecycle. Any ML project expected to have a longer lifespan with multiple updates should aim for MLOps Level 2. Unfortunately, implementing MLOps Level 2 requires an infrastructure that might not be ready – e.g., when models must be approved by different departments that have not yet automated their auditing processes. In those cases, MLOps Level 1 might be the only option for the project's early stages.

The idiosyncrasies of defining and automating ML processes at large-scale organizations are accurately illustrated in the case study by Haakman et al. (2020). While the benefits of having CI/CD and continuous learning enabled are clear, there are a number of legacy processes that do not naturally fit a transition to ML. The case study conducted on a large fintech organization with over 50K employees shows that feasibility assessments, documentation, model risk assessment, and model monitoring are stages where they are far from enabling mature MLOps processes (Haakman et al., 2020). Moreover, the case study also emphasizes the importance of education to get the full potential of ML in practice (Heck & Schouten, 2021).

MLOps Engineering Practice

MLOps is an ever-changing field, and applying it in practice is not yet straightforward. The best way to apply MLOps effectively is by learning from the experiences of other practitioners. Previous works have leveraged catalogs of best practices based on the author's experiences (Akkiraju et al., 2020; Breuel, 2020; Google Cloud, 2020; Koller, 2020; Valohai et al., 2020; Visengeriyeva et al., n.d.). According to these works, the following MLOps best practices ought to be considered when developing an ML system:

- Automate data validation – offline and online.
- Version control everything: data, code, models, parameters, and environment.
- Keep track of experiments via automation. Record information about each experiment to help with data and artefacts lineage, reproducibility, and comparisons.
- Design the right metrics to track and set the threshold of their acceptable values, usually empirically, and often by comparing with previous models or benchmarks.
- Use pipelines and automation.
- Implement automated testing. Test code and models continuously throughout the lifecycle and update test sets accordingly.
- Implement CI/CD. Every training pipeline needs to produce a deployable artefact, not "just" a model.
- Continuous Training (CT). Automate the ML production pipelines to retrain the models with new data.
- Continuous Monitoring (CM). Actively monitor the quality of your model in production.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

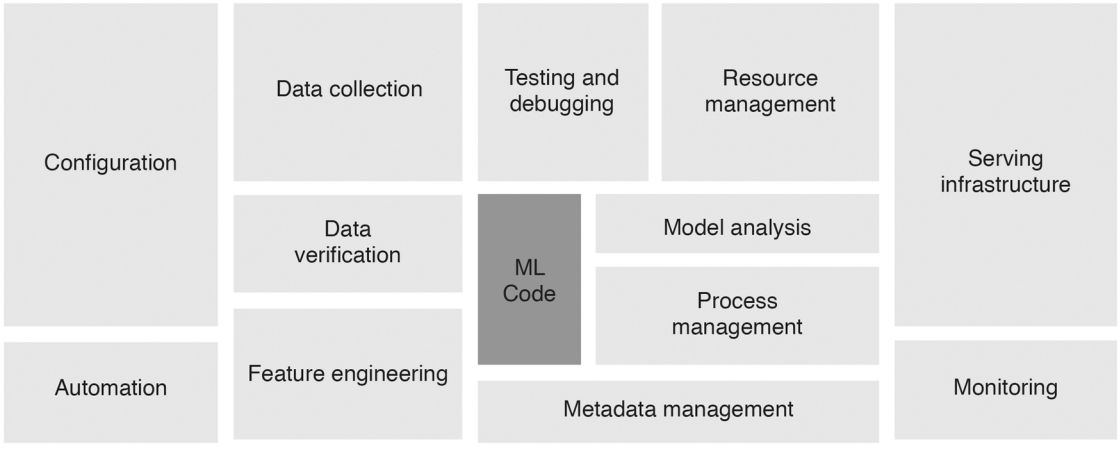
- Work in hybrid teams, including data science, data engineering, ML engineering, and DevOps.
- Write readable code.
- Document data, model, code, and feature dependencies.

These best practices are mostly consensual across different authors, which shows that similar challenges arise in different organizations trying to build production-ready ML systems. Some works take a step further and boil these practices down to a set of principles (Visengeriyeva et al., 2020) or define a manifesto of MLOps for "reproducible, accountable, collaborative, and continuous" ML systems (mlops.community/manifesto/). Such a unified view on the full ML lifecycle, including the surrounding software, is quintessential to build successful production-ready ML systems.

MLOps Supporting Tools

As demonstrated earlier in this chapter, the ultimate goal of MLOps is to automate the complete ML workflow, version and monitor it, such that ML systems can be continuously improved in a reproducible and accountable manner. At the same time, different roles (data scientist, data engineer, ML engineer, operations engineer) should work together throughout the whole lifecycle. Figure 6 shows that a large supporting infrastructure is required for an ML system (Google Cloud, 2020), which is only achievable with considerable tool support. There are many choices for tool support, ranging from using an off-the-shelf machine-learning platform to implementing an in-house solution by composing open-source libraries (Visengeriyeva et al., n.d.).

Figure 6. Only a small fraction of a real-world ML system is composed of the ML code (adapted from Google Cloud, 2020)



Since MLOps technologies are rapidly changing, it is difficult to select the right tool or platform for a given task or problem. There are a number of resources that can help to get a broad picture of all the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

different technologies. IT-consultancy organization Gartner publishes the yearly *Magic Quadrant for Data Science and ML Platforms* (Krensky et al., 2020). The report analyses existing platforms that support MLOps and compare them through a scoring system called the Gartner Magic Quadrant. However, they only score vendors with commercially licensable products.

As an alternative to Gartner's list, the Institute for Ethical AI & Machine Learning created the open-source list "Awesome production ML" (Github, 2020). As part of the list, a taxonomy of the different categories of tools is leveraged, as depicted in Table 1. For completeness, the repository includes commercial platforms in a separate category.

Table 1. Categories of open-source tool support for production ML, adapted from (Github, 2020)

Explaining predictions & models	Privacy preserving ML	Model & data versioning
Model training orchestration	Model serving and monitoring	Neural architecture search
Reproducible notebooks	Visualization frameworks	Industry-strength NLP
Data pipelines & ETL	Data labelling	Data storage
Functions as a service	Computation distribution	Model serialization
Optimized calculation frameworks	Data stream processing	Outlier and anomaly detection
Feature engineering	Feature stores	Adversarial robustness

New tool support is coming out almost every day. The take-home message is that designing and automating an ML pipeline is an activity by its own. It needs careful consideration of each activity's desired level of automation, the choice for cloud or on premise, and if one single platform is preferable over best-of-breed components. Having tailored MLOps is mandatory for building production-ready ML systems. The MLOps platform itself should also be designed with continuous improvement (when tools change or usage scenarios change) in mind.

RELATED WORK

Only recently, research started to focus more on the complete lifecycle of the ML project. Both Marbán et al. (2009) and Hesenius et al. (2019) realize that software engineering principles could help manage the ML lifecycle. Ozkaya (2020) adds that software engineering has solutions "to deal with uncertainty in building, deploying, and sustaining software systems". Dealing with uncertainty is paramount for building production-ready ML systems. This chapter follows a similar line of thinking by identifying challenges that differentiate ML systems from software systems and presenting software engineering solutions for those challenges. The challenges and solutions in this chapter are presented as a unified list but came from diverse sources.

Challenges for Building ML Systems

Amershi et al. (2019) conducted a case study at Microsoft. They interviewed 14 software engineers, mainly in senior leadership roles. Based on the results of the interviews, they surveyed another 551 soft-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ware engineers. They found that various Microsoft teams have united the ML workflow into preexisting, well-evolved, Agile-like software engineering processes. The case study provides insights into several essential engineering challenges that organizations may face in creating large-scale ML solutions for the marketplace. They also collected some best practices from Microsoft teams to address these challenges.

Belani et al. (2019) identified challenges specific to requirements engineering (RE) for ML-based systems. They present their challenges using a taxonomy, in which each of the challenges has been mapped to an ML-related entity (model, data, system), as well as a given RE activity. Lwakatare et al. (2019) present a taxonomy of challenges for all software engineering aspects, mapping to evolution stages and model lifecycle activities. The same authors (Lwakatare et al., 2020) also published a systematic literature review where they identified 72 papers related to the development and maintenance of large-scale ML systems in industrial settings. They extracted 23 challenges and eight solutions from those papers. Arpteg et al. (2018) studied seven projects building deep learning systems to identify twelve main associated challenges.

Other authors have conducted surveys to identify challenges from professionals in practice. Ishikawa and Yoshioka (2019) administered a questionnaire to 278 people who have worked on ML-based systems to identify the essential difficulties and their causes. Kim et al. (2017) conducted a large-scale survey with 793 professional data scientists at Microsoft. They discuss the challenges that they face and the best practices they share with other data scientists. T. Zhang et al. (2019) present a large-scale empirical study of deep learning questions in a popular Q&A website, Stack Overflow. They used a list of 39,628 deep learning questions to identify what challenges developers commonly face when building deep learning applications.

This chapter distinguishes five main challenges for building ML systems based on the work by authors mentioned above and other papers discussing individual challenges. The chapter focuses on the challenges for which engineering solutions already exist because the goal of the chapter is to present those solutions.

Best Practices for Building ML Systems

While several papers discuss challenges for building ML systems, the list of papers discussing solutions to these challenges is much shorter. Some of the papers mentioned under the challenges section also identified solutions (Amershi et al., 2019; Kim et al., 2017; Lwakatare et al., 2020). De Souza Nascimento et al. (2019) propose two checklists as a solution to three specific challenges (identifying the clients' business metrics, lack of a defined development process, and designing the database structure), as identified in interviews with seven ML developers from three small software companies. Zinkevich (2017) describes what he calls "best practices in ML from around Google" in a list of over 40 rules. Serban et al. (2020) mined both academic and grey literature and identified 29 engineering best practices for ML applications. They surveyed 313 practitioners to determine the degree of adoption for these practices and to validate their perceived effects.

This chapter contains a comprehensive overview of best practices, both from the above papers as well as from papers or grey literature discussing single solutions. The chapter organizes those best practices according to the challenge the best practice is addressing.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE DIRECTIONS

This section discusses future directions for industry and research.

Industry

A trend in software engineering is the emergence of *Software as a Service* (SaaS). This means that the software is hosted online as a service, and organizations can use it without having the software development expertise to develop it themselves. This trend is also noticeable in ML and will continue to grow. Gartner yearly analyzes what they call *Cloud AI Developer Services* (Baker et al., 2020). These are cloud-hosted services/models that allow development teams to leverage ML models via APIs without requiring deep data science expertise. These hosted models deliver services with capabilities in language, vision, and automated ML. Gartner (Baker et al., 2020) states that "this market is far from mature, however, with new services and capabilities being introduced rapidly. This trend will continue, such that gaps in major vendors' service portfolios are likely to be filled in the next few years. In addition, we expect many vendors to apply AI and ML services to specific use cases and to offer these services to developers. All this will make for a confusing and volatile market, especially as AI continues to advance and mature."

This last remark is not only true for Cloud AI Services, but for the entire market (commercial and open-source) in ML tool support. In the next few years, the market will hopefully mature and consolidate to enable all organizations to build production-ready ML systems with ease.

For using the before-mentioned Cloud AI Developer Services, one still has to have experience as software developer, know how to call APIs, and process the answers. There are many application domains (health, business process improvement, factory automation, etc.) that could benefit from ML, where the solution designers are neither data scientists nor software engineers. Yet, more and more domain-oriented platforms include ML modules for use by domain experts. They might not even know they are using ML under the hood. Examples of these are low-code platforms like Mendix (mendix.com) and Appian (apian.com) and Robotic Process Automation (RPA) tools like UIPath (uipath.com) and Automation Anywhere (automationanywhere.com).

The use of this integrated tooling by non-ML experts is also a solution for another issue that arises in the industry: the number of ML applications is increasing more rapidly than the number of skilled people to develop them. According to Karnowski and Fadely (2016), a new role is emerging: the AI professional.

Figure 7 shows that for data scientists to become builders of intelligent systems (ML systems), they need to skill themselves in software engineering and data engineering (and vice versa). This is a challenge for education (Heck & Schouten, 2021; Menzies, 2019).

Research

Ozkaya (2020) claims that although ML engineering has specific challenges, these challenges are not new. She argues that the rush to deploy ML-enabled systems increases the urgency of making progress on how to model, analyze, and safeguard against our systems' inherent uncertainty. At the same time, she recognizes that research is needed on development process models for ML engineering to incorporate data scientists.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. A new role for the AI professional, adapted from (Karnowski & Fadely, 2016)

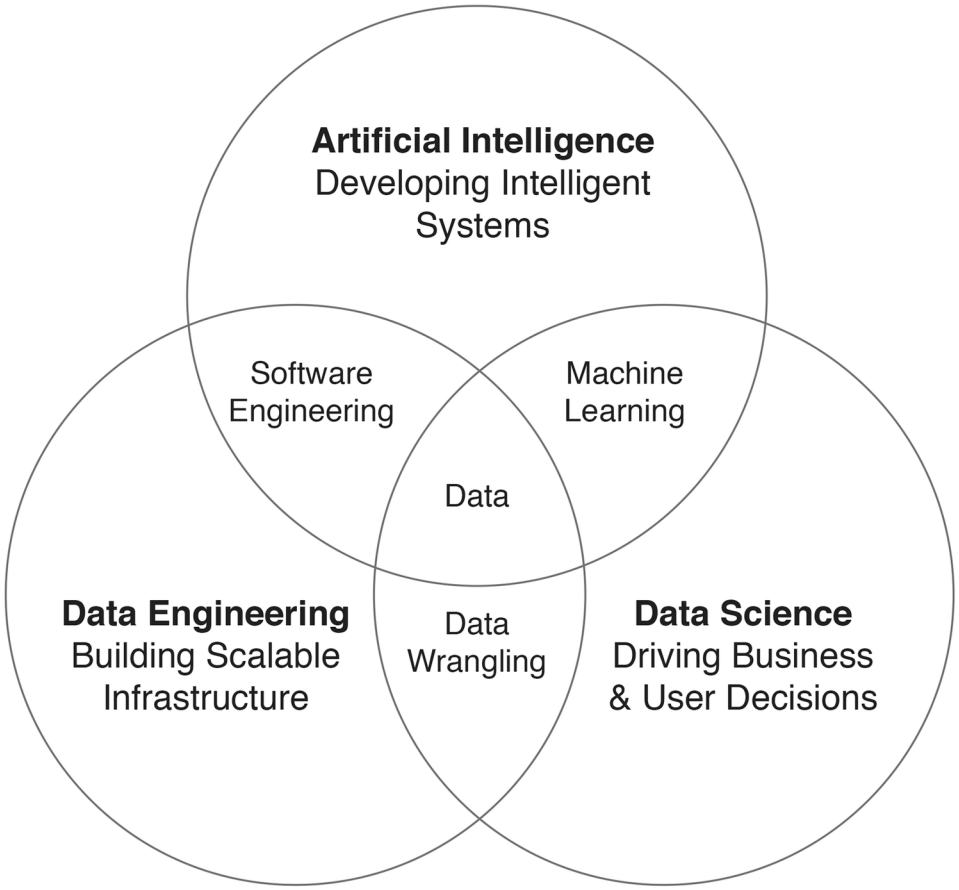


Table 2. A research agenda for ML engineering, adapted from (Bosch et al., 2020)

Strategic Focus	Data Science	ML Engineering		
		Architecture	**Development**	**Process**
data quality management	**Assemble dataset**	Data versioning & dependency mgmt.	DataOps	(automated) labeling
design methods and processes	**create/evolve ML model**	federated learning infrastructure	re-use of pre-developed models	manage multiple models
model performance	**train & evaluate**	storage and computing infrastructure	quality attributes	A/B testing of models
deployment & compliance	**deploy**	deployment infrastructure	integration of models & components	monitoring & logging

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ashmore et al. (2019) provide an overview of research directions in the area of assuring the ML lifecycle, especially when more safety-critical systems have to be ML-enabled. They conclude that "considerable research is still needed to address outstanding assurance concerns associated with every stage of the ML lifecycle. In general, using ML components within safety-critical systems poses numerous open challenges. At the same time, the research required to address these challenges can build on a promising combination of rigorous methods developed by several decades of sustained advances in ML, in software and systems engineering, and in assurance development."

Bosch et al. (2020) provide a comprehensive research agenda for ML engineering. Table 2 contains the domain-independent part of the research agenda. As can be seen, these are similar to the challenges discussed in this chapter and in line with the emergence of an MLOps movement. The MLOps movement in the industry should be fed by new process models, techniques, technologies, and supporting tools coming from academic research.

Note that most references cited in this chapter qualify as so-called grey literature (Garousi et al., 2016). Even more so, several fundamental non-grey publications actually come from the research departments of leading data and ML organizations, not being academics. This might indicate that in the area of ML engineering, research is currently lagging behind industry practice. Thus, there is a need to further establish the research community for ML engineering together with (research) organizations from practice.

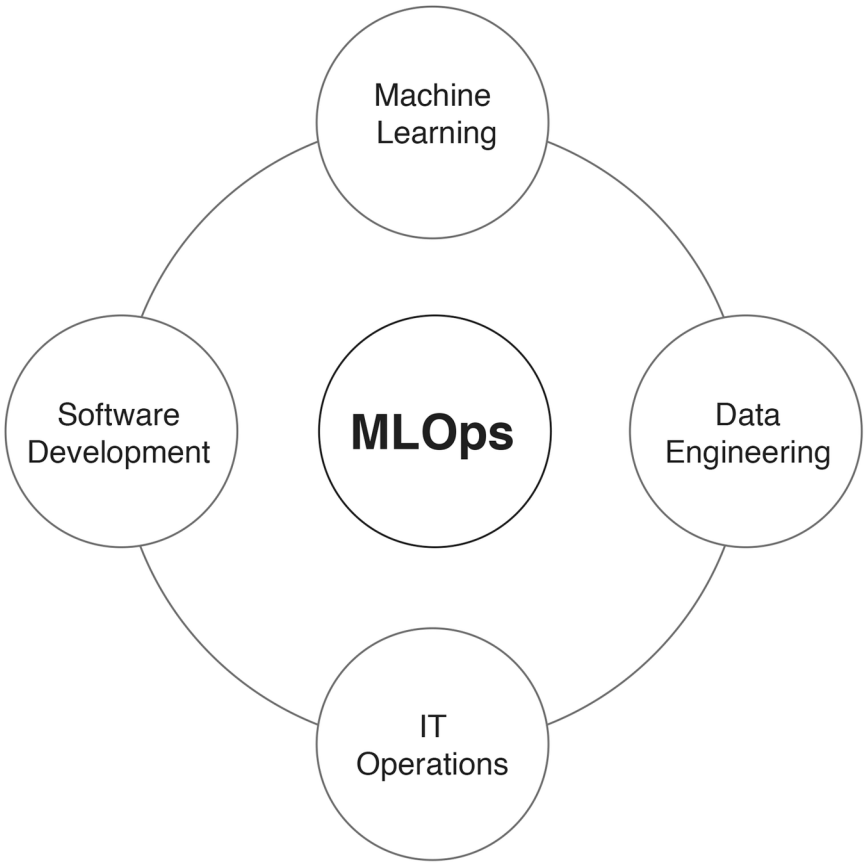
CONCLUSION

This chapter has given a comprehensive overview of what it takes to build production-ready ML systems. In more detail, this chapter has:

1. introduced the term ML engineering and lifecycle stages for ML engineering;
2. discussed the five most important challenges for ML engineering:
   [CH1] elicitation of data and model requirements,
   [CH2] modularizing the application,
   [CH3] design through experimentation,
   [CH4] data and model management,
   [CH5] testing;
3. listed existing engineering solutions for each of the individual challenges;
4. introduced MLOps (see Figure 8) as an overarching and integrated approach to address the complete set of challenges, with the following main principles:
   a. the ML system is developed with a collaborative team across the full ML lifecycle from data to deployment;
   b. the ML system delivers reproducible and traceable results;
   c. the ML system is continuously monitored and improved;
5. given an overview of future developments in industry and research.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The MLOps approach unifies data scientists, software-, data- and operations engineers



MLOps and the individual solutions have only recently emerged in the industry, so the status of solutions discussed in this chapter is volatile. Tools supporting the MLOps approach are likely to evolve and mature over time, as the three main principles mentioned above can only be reached with considerable automation. Academic research into ML engineering will lead to new insights, technologies, methods, and tools for MLOps. These research advancements must be in line with the needs of the industry. Hence, we need case studies with industry applications of ML engineering to learn what works in practice.

REFERENCES

Abbasi-Sureshjani, S., Raumanns, R., Michels, B. E. J., Schouten, G., & Cheplygina, V. (2020). Risk of training diagnostic algorithms on data with demographic bias. *Lecture Notes in Computer Science*, *12446*, 183–192. doi:10.1007/978-3-030-61166-8_20

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access: Practical Innovations, Open Solutions*, *6*, 52138–52160. doi:10.1109/ACCESS.2018.2870052

Akkiraju, R., Sinha, V., Xu, A., Mahmud, J., Gundecha, P., Liu, Z., Liu, X., & Schumacher, J. (2020). *Characterizing machine learning processes: A maturity framework*. Paper presented at the International Conference on Business Process Management. 10.1007/978-3-030-58666-9_2

Algorithmia. (2020). *The roadmap to machine learning maturity*. Retrieved from https://info.algorithmia.com/hubfs/2018/Whitepapers/The_Roadmap_to_Machine_Learning_Maturity_Final.pdf

Alliance for Qualification. (2019). *AI and Software Testing Foundation Syllabus*. Retrieved from https://www.alliance4qualification.info/a4q-ai-and-software-testing

Alspaugh, S., Zokaei, N., Liu, A., Jin, C., & Hearst, M. A. (2018). Futzing and moseying: Interviews with professional data analysts on exploration practices. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 22–31. doi:10.1109/TVCG.2018.2865040 PMID:30136976

Ameisen, E. (2018). *How to deliver on Machine Learning projects. A guide to the ML Engineering Loop.* Retrieved from https://blog.insightdatascience.com/how-to-deliver-on-machine-learning-projects-c8d82ce642b0

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019). *Software engineering for machine learning: A case study.* Paper presented at the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). 10.1109/ICSE-SEIP.2019.00042

Arpteg, A., Brinne, B., Crnkovic-Friis, L., & Bosch, J. (2018). *Software engineering challenges of deep learning.* Paper presented at the 2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). 10.1109/SEAA.2018.00018

Ashmore, R., Calinescu, R., & Paterson, C. (2019). *Assuring the machine learning lifecycle: Desiderata, methods, and challenges.* arXiv preprint arXiv:1905.04223.

Bailis, P., Olukotun, K., Ré, C., & Zaharia, M. (2017). *Infrastructure for usable machine learning: The stanford dawn project.* arXiv preprint arXiv:1705.07538.

Baker, V., Elliot, B., Sicular, S., Mullen, A., & Brethenoux, E. (2020). *Magic quadrant for cloud AI developer services*. Retrieved from https://www.gartner.com/en/documents/3981253/magic-quadrant-for-cloud-ai-developer-services

Belani, H., Vukovic, M., & Car, Ž. (2019). *Requirements engineering challenges in building AI-based complex systems.* Paper presented at the 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW). 10.1109/REW.2019.00051

Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., & Mojsilović, A. (2019). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development, 63*(4/5), 1-4.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Benton, W. C. (2020). Machine learning systems and intelligent applications. *IEEE Software*, *37*(04), 43–49. doi:10.1109/MS.2020.2985224

Bird, S., Dudík, M., Edgar, R., Horn, B., Lutz, R., Milan, V., Sameki, M., Wallach, H., & Walker, K. (2020). *Fairlearn: A toolkit for assessing and improving fairness in AI*. Retrieved from https://www.microsoft.com/en-us/research/uploads/prod/2020/05/Fairlearn_WhitePaper-2020-09-22.pdf

Bosch, J., Crnkovic, I., & Olsson, H. H. (2020). *Engineering AI systems: A research agenda.* arXiv preprint arXiv:2001.07522.

Bourque, P., & Fairley, R. (2014). Guide to the software engineering body of knowledge (SWEBOK (R)): Version 3.0. IEEE Computer Society Press.

Bozonier, J. (2015). *Test-driven machine learning*. Packt Publishing Ltd.

Breck, E., Cai, S., Nielsen, E., Salib, M., & Sculley, D. (2016). *What's your ML Test Score? A rubric for ML production systems*. Paper presented at the Reliable Machine Learning in the Wild - NIPS 2016 Workshop.

Breuel, C. (2020, Jan 3). *MLOps: Machine learning as an engineering discipline*. Retrieved from https://towardsdatascience.com/ml-ops-machine-learning-as-an-engineering-discipline-b86ca4874a3f

Burkov, A. (2020). *Machine learning engineering*. True Positive Inc.

Cai, S., Bileschi, S., Nielsen, E. D., & Chollet, F. (2020). *Deep learning with JavaScript*. Manning.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. Academic Press.

Chen, C., Härdle, W. K., & Unwin, A. (2007). *Handbook of data visualization*. Springer Science & Business Media.

Cliffydcw (Producer). (2020). *Own work*. Retrieved from https://commons.wikimedia.org/w/index.php?curid=19054763

De Souza Nascimento, E., Ahmed, I., Oliveira, E., Palheta, M. P., Steinmacher, I., & Conte, T. (2019). *Understanding development process of machine learning systems: Challenges and solutions.* Paper presented at the 2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). 10.1109/ESEM.2019.8870157

Ebert, C., Gallardo, G., Hernantes, J., & Serrano, N. (2016). DevOps. *IEEE Software*, *33*(3), 94–100. doi:10.1109/MS.2016.68

Farah, D. (2020). *The modern MLOps blueprint*. Retrieved from https://medium.com/slalom-data-analytics/the-modern-mlops-blueprint-c8322af69d21

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Garousi, V., Felderer, M., & Mäntylä, M. V. (2016). *The need for multivocal literature reviews in software engineering: Complementing systematic literature reviews with grey literature.* Paper presented at the 20th International Conference on Evaluation and Assessment in Software Engineering. 10.1145/2915970.2916008

Github. (2020). *Awesome production machine learning*. Retrieved from https://github.com/EthicalML/awesome-production-machine-learning

Google Cloud. (2020). *MLOps: Continuous delivery and automation pipelines in machine learning*. Retrieved from https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

Haakman, M., Cruz, L., Huijgens, H., & van Deursen, A. (2020). *AI lifecycle models need to be revised. An exploratory study in fintech.* arXiv preprint arXiv:2010.02716.

Heck, P. (2019). *Software engineering for machine learning applications*. Retrieved from https://fontysblogt.nl/software-engineering-for-machine-learning-applications/

Heck, P. (2020). *Testing machine learning applications*. Retrieved from https://fontysblogt.nl/testing-machine-learning-applications/

Heck, P., & Schouten, G. (2021). *Lessons learned from educating AI engineers*. Paper accepted for the 1st International Workshop on AI Engineering (WAIN).

Hesenius, M., Schwenzfeier, N., Meyer, O., Koop, W., & Gruhn, V. (2019). *Towards a software engineering process for developing data-driven applications.* Paper presented at the 2019 IEEE/ACM 7th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE). 10.1109/RAISE.2019.00014

Hulten, G. (2019). *Building intelligent systems*. Apress. doi:10.1007/978-1-4842-3933-9

Institute for Ethical AI & Machine Learning. (2020). *Machine learning maturity model v1.0*. Retrieved from https://ethical.institute/mlmm.html

Ishikawa, F., & Yoshioka, N. (2019). *How do engineers perceive difficulties in engineering of machine-learning systems? Questionnaire survey.* Paper presented at the 2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SER&IP). 10.1109/CESSER-IP.2019.00009

Jurney, R. (2017). *Agile data science 2.0: Building full-stack data analytics applications with Spark*. O'Reilly Media.

Kang, H., Le, M., & Tao, S. (2016). *Container and microservice driven design for cloud infrastructure devops.* Paper presented at the 2016 IEEE International Conference on Cloud Engineering (IC2E). 10.1109/IC2E.2016.26

Karnowski, J., & Fadely, R. (2016). *How AI careers fit into the data landscape*. Retrieved from https://blog.insightdatascience.com/how-emerging-ai-roles-fit-in-the-data-landscape-d4cd922c389b

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010). *Mastering the information age: Solving problems with visual analytics*. VisMaster.

Kery, M. B., Radensky, M., Arya, M., John, B. E., & Myers, B. A. (2018). *The story in the notebook: Exploratory data science using a literate programming tool.* Paper presented at the CHI Conference on Human Factors in Computing Systems. 10.1145/3173574.3173748

Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2017). Data scientists in software teams: State of the art and challenges. *IEEE Transactions on Software Engineering*, *44*(11), 1024–1038. doi:10.1109/TSE.2017.2754374

Kirk, M. (2017). *Thoughtful machine learning with Python: A test-driven approach*. O'Reilly Media.

Koller, B. (2020, Sep 28). *12 Factors of reproducible Machine Learning in production*. Retrieved from https://blog.maiot.io/12-factors-of-ml-in-production/

Krensky, P., den Hamer, P., Brethenoux, E., Hare, J., Idoine, C., Linden, A., Sicular, S., & Choudhary, F. (2020). *Magic quadrant for data science and machine learning platforms*. Academic Press.

Kriens, P., & Verbelen, T. (2019). *Software engineering practices for machine learning.* arXiv preprint arXiv:1906.10366.

Kühn, A., Joppen, R., Reinhart, F., Röltgen, D., von Enzberg, S., & Dumitrescu, R. (2018). Analytics Canvas–A Framework for the design and specification of data analytics projects. *Procedia CIRP*, *70*, 162–167. doi:10.1016/j.procir.2018.02.031

Lwakatare, L. E., Raj, A., Bosch, J., Olsson, H. H., & Crnkovic, I. (2019). *A taxonomy of software engineering challenges for machine learning systems: An empirical investigation.* Paper presented at the International Conference on Agile Software Development. 10.1007/978-3-030-19034-7_14

Lwakatare, L. E., Raj, A., Crnkovic, I., Bosch, J., & Olsson, H. H. (2020). Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions. *Information and Software Technology*, *127*, 106368. doi:10.1016/j.infsof.2020.106368

Marbán, O., Segovia, J., Menasalvas, E., & Fernández-Baizán, C. (2009). Toward data mining engineering: A software engineering approach. *Information Systems*, *34*(1), 87–107. doi:10.1016/j.is.2008.04.003

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Vintage.

McAteer, M. (2020). *Nitpicking machine learning technical debt*. Retrieved from https://matthewmcateer.me/blog/machine-learning-technical-debt/

Menzies, T. (2019). The five laws of SE for AI. *IEEE Software*, *37*(1), 81–85. doi:10.1109/MS.2019.2954841

Microsoft. (2020a). *What are Azure Machine Learning pipelines?* Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/concept-ml-pipelines

Microsoft. (2020b). *What is the Team Data Science Process?* Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/overview

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). *Model cards for model reporting.* Paper presented at the Conference on Fairness, Accountability, and Transparency. 10.1145/3287560.3287596

Molnar, C. (2020). *Interpretable machine learning*. Lulu.Com.

Nori, H., Jenkins, S., Koch, P., & Caruana, R. (2019). *Interpretml: A unified framework for machine learning interpretability.* arXiv preprint arXiv:1909.09223.

Oved, D. (2018). *Real-time human pose estimation in the browser with TensorFlow.js*. Retrieved from https://medium.com/tensorflow/real-time-human-pose-estimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5

Ozkaya, I. (2020). What is really different in engineering AI-enabled systems? *IEEE Software*, *37*(4), 3–6. doi:10.1109/MS.2020.2993662

Pinhasi, A. (2020). *Deploying machine learning models to production: Inference service architecture patterns*. Retrieved from https://medium.com/data-for-ai/deploying-machine-learning-models-to-production-inference-service-architecture-patterns-bc8051f70080

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

Raman, A., & Hoder, C. (2020). *Building intelligent apps with cognitive APIs*. O'Reilly.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). *"Why should I trust you?": Explaining the predictions of any classifier.* Paper presented at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA. 10.1145/2939672.2939778

Rollins, J. (2015). *Why we need a methodology for data science*. Retrieved from https://www.ibmbigdatahub.com/blog/why-we-need-methodology-data-science

Sato, D., Wider, A., & Windheuser, C. (2019). *Continuous delivery for machine learning: Automating the end-to-end lifecycle of machine learning applications*. Retrieved from https://martinfowler.com/articles/cd4ml.html

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., & Young, M. (2014). *Machine learning: The high interest credit card of technical debt*. Paper presented at the SE4ML Workshop (NIPS 2014).

Serban, A., van der Blom, K., Hoos, H., & Visser, J. (2020). *Adoption and effects of software engineering best practices in machine learning*. Paper presented at the ESEM conference. 10.1145/3382494.3410681

Sommerville, I. (2019). *Artificial intelligence and systems engineering*. Academic Press.

Sommerville, I., & Sawyer, P. (1997). *Requirements engineering: A good practice guide*. John Wiley & Sons, Inc.

Spiekermann, S. (2015). *Ethical IT innovation: A value-based system design approach*. CRC Press. doi:10.1201/b19060

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tan, D. (2019). *Coding habits for data scientists*. Retrieved from https://www.thoughtworks.com/insights/blog/coding-habits-data-scientists

Thalheim, B. (2013). *Entity-relationship modeling: Foundations of database technology*. Springer Science & Business Media.

Valohai, SigOpt, & Tecton. (2020). *Practical MLOps: How to get ready for production models*. Author.

Visengeriyeva, L., Kammer, A., Bär, I., Kniesz, A., & Plöd, M. (n.d.). *Machine learning operations*. Retrieved from https://ml-ops.org/

Vogelsang, A., & Borg, M. (2019). *Requirements engineering for machine learning: Perspectives from data scientists.* Paper presented at the IEEE 27th International Requirements Engineering Conference Workshops (REW). 10.1109/REW.2019.00050

Wang, D., Weisz, J. D., Muller, M., Ram, P., Geyer, W., Dugan, C., Tausczik, Y., Samulowitz, H., & Gray, A. (2019). Human-AI collaboration in data science: Exploring data scientists' perceptions of automated AI. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1-24. 10.1145/3359313

Washizaki, H., Uchida, H., Khomh, F., & Guéhéneuc, Y.-G. (2020). *Machine learning architecture and design patterns*. Academic Press.

Yan, E. (2019). *Data science and agile: What works and what doesn't work*. Retrieved from https://towardsdatascience.com/agiledatascience-3b7ca65278a4

Zhang, J., Harman, M., Ma, L., & Liu, Y. (2020). *Machine learning testing: Survey, landscapes and horizons. IEEE Transactions on Software Engineering*. Early Access.

Zhang, T., Gao, C., Ma, L., Lyu, M., & Kim, M. (2019). *An empirical study of common challenges in developing deep learning applications.* Paper presented at the 2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE). 10.1109/ISSRE.2019.00020

Zinkevich, M. (2017). *Rules of machine learning: Best practices for ML engineering*. Retrieved from https://developers.google.com/machine-learning/guides/rules-of-ml

ADDITIONAL READING

Dubovikov, K. (2019). *Managing data science*. Packt.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.

Github. (2020a). Awesome MLOps. Retrieved from https://github.com/visenger/awesome-mlops

Kaestner, C. (2020). Software engineering for AI/ML: An annotated bibliography. Retrieved from https://github.com/ckaestne/seaibib

Lakshmanan, V., Robinson, S., & Munn, M. (2020). *Machine learning design patterns*. O'Reilly Media.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ng, A. (2019). *Machine learning yearning: Technical strategy for AI engineers in the era of deep learning*. Retrieved from https://www.deeplearning.ai/machine-learning-yearning/

Patruno, L. (2020). The ultimate guide to deploying machine learning models. Retrieved from https://mlinproduction.com/deploying-machine-learning-models/

Smith, J. (2018). *Machine learning systems: Designs that scale*. Manning Publications Co.

Sweenor, D., Hillion, S., Rope, D., Kannabiran, D., Hill, T., & O'Connell, M. (2020). *MLOps: Operationalizing data science*. O'Reilly.

Taifi, M. (2020). *Clean machine learning code*. Leanpub.

KEY TERMS AND DEFINITIONS

CI/CD: Continuous Integration and Continuous Deployment, a set of practices to automate testing, integration, and deployment into a production environment of software components.

CT: Continuous Training, an extension of CI/CD for machine learning systems by automating the retraining and serving in a production environment of machine learning models.

DevOps: A set of best practices that aims at unifying software system development (Dev) and software system operation (Ops).

ML Engineering: The discipline of (software) engineering applied to the development, operation, and maintenance of machine learning (ML) systems.

ML System: A software system that contains a machine learning (ML) component.

MLOps: A set of best practices that aims at unifying machine learning (ML) system development and machine learning system operation (Ops).

Pipeline: In software engineering, a pipeline consists of a chain of processing elements arranged so that the output of each element is the input of the next; the name is by analogy to a physical pipeline.

Production Environment: A term used in software development to describe the setting where software products are put into operation and made available for end users.

Software Engineering: A computing discipline that advocates the application of engineering principles to the development, operation, and maintenance of software.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 3
Multivariate Sustainability Profile of Global Fortune 500 Companies Using GRI-G4 Database

Mónica Jiménez-Hernández
Universidad de Colima, Mexico

Purificación Vicente-Galindo
University of Salamanca, Spain

Nathalia Tejedor-Flores
Universidad Tecnológica de Panamá, Panama

Adelaide Freitas
University of Aveiro, Portugal

Purificación Galindo
University of Salamanca, Spain

ABSTRACT

The main objective of this research is to find the sustainability gradients of Global Fortune 500 companies and sort them as a function of economic, environmental, and social components using multivariate statistical methods to establish the foundations for better knowledge of the trends and sustainability reporting habits. A combined approach, comprising principal coordinates analysis (PCoA) and logistic regression model (LRM), is proposed to build an external logistics biplot (ELB). Moreover, HJ-Biplot and parallel coordinates are applied. This chapter helps to understand why many companies view their corporate social responsibility (CSR) reports as a way to guarantee the credibility of the published information. In particular, based on the Global Reporting Initiative, the sustainability gradients of the Global Fortune 500 companies are obtained and statistically exploited to analyze how the companies can make improvements in terms of sustainability.

DOI: 10.4018/978-1-7998-6985-6.ch003

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The growing and constant increase in the population and the search to satisfy their needs without compromising the environment and society, have been leading to the creation of the concept of *sustainable development*, which beyond just focusing on environmental matters, integrates economic and social aspects. Nowadays, it would be difficult to think of societies where companies are just entities exclusively oriented to the generation of profits for their shareholders. Corporate Social Responsibility (CSR) has gradually become relevant along with the creation of economic value (Tejedor-Flores, Galindo-Villardón, & Vicente-Galindo, 2017).

A sustainable company is one that creates economic, social, and environmental value in the short and long term, hence contributing to the increase in the welfare of society and the progress of present and future generations. Sustainable development in a company is essential to maintain the economic profitability of its activities, taking into account new concepts of risk and opportunity. These are associated with environmental aspects and the social impact of production or the quality of labour relations (Rodriguez Rojas, 2015). The demand for information and transparency from companies has increased over the last few years. Public awareness and interest in social, environmental, and ethical issues and increased attention from the mass media have resulted in more companies making social disclosures about themselves (Ho & Taylor, 2007; Rouf, 2011; Patten, 1991). To defend a company's sustainability management, it is of the main importance that it shows this in its sustainability reports. The management of the company should try to meet the expectations of all stakeholders (investors, consumers, government, among others). This is why companies produce this type of publication, to be transparent and make themselves known in terms of sustainability.

To satisfy these expectations and clearly and openly inform about sustainability, the Global Reporting Initiative (GRI) was founded in 1997. Its roots lay in the US non-profit organizations the Coalition for Environmentally Responsible Economies (CERES) and the Tellus Institute. GRI is an international independent organization that helps businesses, governments, and other organizations understand and communicate the impact of business on critical sustainability issues such as climate change, human rights, corruption, and many others (Global Reporting Initiative, 2013).

This chapter addresses the G4 Guideline reports corresponding to the four-generation of the GRI. This Guideline envision a sustainable global economy that combines long-term profitability with social justice and environmental care. The G4 Guidelines are organized into three categories: Economic, Environment, and Social. The Social category is extended into the following subcategories: labour, human rights, society, and product responsibility (Global Reporting Initiative, 2013). This chapter attempts to analyze, using multivariate statistical techniques, a group of companies with the highest income worldwide to find out their reporting characteristics. Using the GRI-G4 database, questions as following are study targets: What are the most relevant and frequently reported indicators? How to identifying sustainability gradients of these companies and similar groups of companies? Besides, these multivariate techniques are allowing the analyis of the sustainable development by sector, region, and country, and consequently, the findings of the sustainability characteristics of the companies corresponding to these areas.

The chapter is structured as follows: Section 1: Theoretical framework, this section presents the main definitions of sustainable development and sustainability, as well as a summary of the development of these concepts. This section also presents a brief history of the Global Reporting Initiative (GRI) as a pioneer in sustainability reporting worldwide. Section 2: Fundamental theory of multivariate statistic, this section contains the fundamental theory on the statistical techniques used, like External Logistics

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Biplot, HJ-Biplot, and Parallel Coordinates. Section 3: Methodology, this section describes each of the analysis techniques used in this work, as well as their interpretation. It also describes the way in which the information was obtained, the creation of the database, and details the sustainable variables studied. Section 4: Results, the main results of the analyses are presented in descriptive statistics, External Logistic Biplot, HJ-Biplot, and Parallel Coordinates. Section 5: Discussion of the results, and finally, Section 6: Conclusions.

Theoretical Framework

Sustainability appears as "the central unifying idea most needed at this moment in human history", according to (Bybee, 1991). It is a concept that seeks to mobilize collective responsibility to face the set of serious problems and challenges that humanity is facing, by betting on cooperation and the defence of the general interest. Nowdays, new scientific knowledges have emerged new possibilities to evaluate and discuss diverse aspects of phenomena that occur around the world and can affect the environmental quality. Phenomena like the existence of the hole in the ozone layer, loss of biodiversity, and global warming. These type of phenomena constitute a link between economic activity and the environment bringing society's concern for better development (Gómez, 2018).

Formally, the concept of sustainable development was introduced in 1987 by Brundtland. This author defined as: "development, which meets the needs of current generations without compromising the ability of future generations to meet their own needs" (Brundtland, 1987). By extending the concept, the United Nations mentions that, in order to achieve sustainable development, it is fundamental to take into account three basic elements that are interrelated and essential for the well-being of both individuals and societies: economic growth, social inclusion, and environment. Although there are currently many definitions of sustainable development, they all agree that, in order to achieve it, action economically viable, environmentally friendly, and socially equitable, must be taken (Miñón, 2002). In the same way, the Institute of Chartered Accountants in England and Wales (ICAEW) mentions that the concept of sustainability implies that organizations should operate taking into account the impacts of sustainability on the planet, the population, and the future (ICAEW, 2004).

Furthermore, the 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart 17 Sustainable Development Goals (SDGs) appear as an urgent call for action by all countries - developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests (United Nations, 2019).

Sustainability Reports

Some authors find the roots of CSR in the 1920s: when the principle of charity and philanthropy, there are no longer as an individual actions (of the entrepreneur or the owners of the company) but a corporate actions of the company as a whole (Davis & Blomstrom, 1980). On the other hand, private business organisations such as the World Business Council for Sustainable Development (WBCSD) define CSR as "the commitment of business to contribute to sustainable economic development by working with

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

employees, their families, the local community and society at large to improve their quality of life" (Holme & Watts, 2000).

Sustainability reports at the corporate level are intended to represent a company's environmental, social, and economic performance, as well as the impacts related to the planet (ICAEW, 2004). For any organization, this type of document provides information on its sustainability performance. Both positive and negative contributions are presented in orderto make a comparison in terms of regulations and laws, and perspectives of sustainable development in the company (Santolaria; & González, 2013). Besides, sustainability reporting helps organizations to set objectives, measure their performance, and manage the change making their operations more sustainable. All these reports contain information on the impact of companies on the environment, society, and the economy (Global Reporting Initiative, 2013).

The Global Reporting Initiative (GRI) is an independent international organization that pioneered sustainability reporting by companies from 1997 to the present day. The GRI helps companies and governments worldwide understand and communicate their impact on sustainability issues, specifically climate change, human rights, governance, and social welfare. The GRI for Sustainability Reporting Standards are developed with genuine multi-stakeholder input and are based on public interest, thus enabling real action to create social, environmental and economic benefits for all (GRI, 2020).

The GRI was founded in Boston, USA, in 1997 and three non-profit organizations were the roots of this project: Coalition for Environmentally Responsible Economies (CERES), the Tellus Institute, and the United Nations Environment Programme (UNEP). A year later, a Management Committee was created and it was established that, rather than focusing on environmental issues, this organization would also advocate for social, economic, and governance issues (Global Reporting Initiative, 2013). Since the GRI issued its first sustainability reporting guidelines in 2000, several have followed (G2, G3, G3.1) with the G4 guideline being the fourth and the version considered in this chapter.

The version of the GRI guidelines used in this chapter is the G4 and like other GRI guidelines, this version refers to various documents on specific aspects of reporting that are widely accepted and used by many companies. This version of the GRI guidelines consists of 58 indicators with general basic content, which are divided into the followingsections: Strategy and analysis, organizational profile, material aspects of coverage, stakeholder engagement, report profile, governance and ethics, and integrity. The indicators used in this work are 91 specific indicators which are partitioned in three categories: Economic, Environmental and Social. These indicators provide information on the company's performance in these three areas (Global Reporting Initiative, 2013). The G4 Guidelines categories:

1. **Economic Category (EC):** This category reflects the flow of capital between the different stakeholders and the economic impacts that the company generates in society at a local, national and international level. This category is made up of 9 indicators that are also classified into four aspects: economic performance, market presence, indirect economic consequences, and acquisitions practices.
2. **Environment Category (EN):** This category covers the impacts on inputs and outputs, among others, for instance, energy consumption, water withdrawal, emissions and waste generated at the time of operation of the company or at the time of providing its services. It also covers aspects of biodiversity, transport, the impact of products or services in the environmental. This category has got 34 indicators that are divided into twelve aspects: materials, energy, water, biodiversity, emissions, effluents and waste, products and services, regulatory compliance, transport, overall, environmental assessment of suppliers, and environmental complaint mechanisms.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. **Social Performance Category:** This category relates to the impact of the company on the social systems where it operates. It is based on a series of internationally recognised standards and is divided into four subcategories:
   a. Labour Practices (LA): this is divided into 8 areas: employment, relations between workers, health and safety at work, training and education, diversity and equal opportunities, equal pay for men and women, evaluation of suppliers labour practices, and mechanisms for complaints about labour practices. The total number of indicators in this subcategory is 16.
   b. Human Rights (HR): there is a total of 12 indicators divided into 10 aspects: investment, non-discrimination, freedom of association and collective bargaining, child labour, forced labour, security measures, indigenous rights, assessment, supplier human rights assessment, and human rights grievance mechanisms.
   c. Society (SO): this is divided into 7 aspects: local communities, the fight against corruption, public policy, unfair competition practices, regulatory compliance, evaluation of the social impact of suppliers, and social impact complaint mechanisms. The total number of indicators in this subcategory is 11.
   d. Product Responsibility (PR): this is related to the products and services that affect the stakeholders. It consists of 9 indicators divided into 5 aspects: customer health and safety, product and service labelling, marketing communications, customer privacy, and regulatory compliance.

In this study, we used the G4 Guidelines, because these guidelines have increased user-friendliness and most of the top 500 companies published by FORTUNE magazine use these guidelines to present their sustainability reports. Also, these guidelines emphasise on what is material encourages organizations to provide only information that is critical to their business and stakeholders. This means organizations and report users can concentrate on the sustainability impacts that matter, resulting in reports that are more strategic, more focused, more credible, and easier for stakeholders to navigate (Buck, Espinach, & Söderberg, 2014).

FUNDAMENTAL THEORY OF MULTIVARIATE STATISTICS

Classical Biplots

The biplot method (Gabriel, 1971) is a simultaneous graphical representation of the rows and the columns of a given data matrix. The main uses are exploratory although it has also been used as a graphical representation for more formal models (Gabriel, 1998). In practice, Biplot fitting occurs either by computing the singular value decomposition (SVD) of the data matrix or by performing an alternating regressions procedure (Gabriel & Zamir, 1979).

A Biplot for a data matrix X ($I \times J$) is a graphical representation using markers (vectors) a_1, a_2,...,a_I for the rows of X and b_1,b_2,...,b_J for the columns of X, such that the inner product approximates the element x_{ij} of the starting matrix as well as possible. If we consider the markers a_1, a_2,...,a_I as rows of a matrix A and the markers b_1, b_2,...,b_J as rows of a matrix B, then we can write: $X \cong AB'$. Both the markers a_i for the rows and the markers b_j for the columns will be represented in a space of dimension $q \leq r$, where q is the number of retained axes and r is the rank of X. The Biplot factorisation guarantees

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the approximate graphical representation of the matrix, since each xij can be reconstructed as follows: $x_{ij} = a_i' b_j \left(i = 1,2,\ldots,I; J = 1,2,\ldots,J\right)$.

Considering the Biplot approximation of the elements of the matrix X and applying the definition of scalar product we have: $x_{ij} = a'_i\, b_j = \left\| Proy(a_i / b_j \left(sign\right) \right\| b_j \|$, where $\left\| Proy(a_i / b_j \right\|$ is the length of the segment from the origin of the representation to the point a_i (length of the projection from a_i to b_j), sign b_j is the sign of bj and b_j is the module of b_j (length of the segment from the origin to b_j). This means that x_{ij} is approximately the module of the projection of a_i onto b_j multiplied by the length of b_j, with its corresponding sign. The direction of the vector b_j shows the increasing direction of the corresponding variable values. The projections of the points a_i onto a column vector approximate the j*th* column of X and provide an ordination of the individuals respect to the corresponding variable. The structure of the X matrix can be visualised by representing the markers in Euclidean space, usually in two or three dimensions.

The objective is to find a matrix $X_{(q)} = A_{(q)}B_{(q)}'$ of rank q, which best approximates X, in the least-squares sense, where $X_{(q)}$ must minimise the following expression: $\sum_i \sum_j \left(x_{ij} - x_{(q)ij}\right)^2 = traza\left[\left(X - X_{(q)}\right)\left(X - X_{(q)}\right)'\right]$ for all matrices X(q) of rank q or less. The best known method for approximating a low-rank matrix is based on the Singular Value Decomposition (SVD) and Singular Vectors of the matrix X (Eckart & Young, 1936): X=UDV', where U (Ixr) is the matrix whose columns contain the eigenvectors of XX'; V (Jxr), the matrix whose columns contain the eigenvectors of X'X and D (rxr) is a diagonal matrix containing the singular values of X.

It must be satisfied that U'U=V'V=I, therefore, the columns of U and V are orthonormal. So the best approximation of X in rank q ($X_{(q)}$) is: $X_{(q)_{Ixp}} = U_{(q)_{Ixq}} D_{(q)_{qxq}} V'_{(q)_{qxJ}} = \sum_{k=1}^{q} \lambda_k u_k v_k'$, where $U_{(q)}$ and $V_{(q)}$, are the matrices constructed from the first q columns of U and V, respectively, and $D_{(q)}$ is the diagonal matrix containing the q largest non-zero singular values of X (λ_k). Therefore: $X = AB' = UDV'$. The selection of different factorisations leads to different types of markers and thus to different Biplots. The interpretation of the Biplot is always based on the scalar products, regardless of the chosen factorisation. The two most important Biplot factorisations proposed by Gabriel (1971) were named: GH-Biplot and JK-Biplot. The GH-Biplot achieves high quality in the representation of the columns (variables), while the JK-Biplot achieves high quality in the representation of the rows (individuals). Therefore, the Biplot representation is not simultaneous for either of these two types of Biplots because the goodness of fit is not the same for the rows and for the columns of the data matrix. A variant, called HJ-Biplot, was proposed by Galindo-Villardón, in 1986.

HJ-BIPLOT

The HJ-Biplot is a multivariate graphical representation of row and column markers, chosen in such a way that they can be overlapped in the same reference system with maximum quality of representation; its main objective is the simultaneous interpretation of the relationships between individuals (rows) and variables (columns) of the data matrix X. In this variant, the individuals are described by the row vectors of the matrix A=UD and displayed as points in the biplot, and the variables are described by the row

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

vectors of the matrix B=VD and displayed as arrows in the biplot. In the HJ-Biplot, the representation of rows and columns of the data matrix is symmetrical, the same quality of representation is obtained for both individuals (rows) and variables (columns).

The rules for interpreting the HJ-Biplot are: the distances between points are interpreted as inverse similarities (closer individuals present more similar row profiles); the lengths of the arrows (vectors) approximate the standard deviation of the correspondent variables; the cosines of the angles between two arrows approximate the correlations between the correspondent variables (Nieto, Galindo-Villardón, Leiva & Vicente-Galindo, 2014).

EXTERNAL LOGISTIC BIPLOT

When data is binary, classical linear biplots are not suitable because the response along the dimensions is linear, in the same way as linear regression is not suitable when response is binary. Multiple correspondence analysis (MCA) is commonly used for categorical data, and can be considered as a particular form of biplot for binary (or indicator) matrices. Gower & Hand (1996) and Gabriel (1995) develop the complete theory for MCA in relation to biplots. Gower & Hand (1996) propose what they call "prediction regions" as an extension of the usual linear projections. The prediction regions are based on distances from the individual points to the category points. The representation space is divided into regions that predict each category or combination of categories.

In this chapter we used a linear biplot for binary data in which response along the dimensions is logistic. Each individual is represented as a point and each character as a direction through the origin. The projection of an individual point onto a character direction predicts the probability of presence of that character. The method is related to logistic regression in the same way that biplot analysis is related to linear regression. This method named as the logistic biplot, was proposed by Vicente–Villardon, Galindo–Villardon & Blazquez–Zaballos in 2006. This algorithm is later extended by an integrated approach as suggested by Demey *et al*. in 2008. This combinatory method comprises a Principal Coordinates Analysis (PCoA) and Logistic Regression Model (LRM) to construct an External Logistic Biplot (ELB).

The algorithm starts with a Principal Coordinate Analysis (PCoA) to order the individuals in a Euclidean space where they are represented as points. (Vicente Galindo, Vaz, & Nijkamp, 2011). The dimensions in this analysis can be interpreted as gradients. The next step is to fit a logistic regression model on the principal coordinate matrix for each indicator using the latent gradients. In this case, the coefficients of the regression are the vectors indicating the direction in which the probability of presence of each indicator is best predicted, resulting in a logistic biplot, named like that, because the coordinates of the individuals are calculated in an external procedure (Demey *et al*., 2008).

In this analysis the quality of representation of the variables must be taken into account for the interpretation of the results, which is summarised in the following three fundamental elements (Cubilla-Montilla *et al*., 2015):

- The p-value to determine the variables that are significant in the analysis.
- The coefficient of determination or R^2.
- The percentage of well-classified variables, obtained from the expected probabilities. Values closer to 100% indicate better classification.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In analytical terms the ELB is interpreted according to the following rules:

- Distances between company points on the scattergram are inversely related to their profile similarities, i.e. companies close together have similar characteristics.
- The angles between vectors representing the variables and factorial axis, estimate the degree of the relation between the variable and the latent dimension. Taking into account that the horizontal axis is the one which always accounts for most information, variables forming acute angles with the first axis are the most relevant to ordering the companies in relation to the more important gradient of sustainability; that is to classify companies according to sustainability.
- The projections of the companies onto the vector representing the variables, estimate the expected probability of the characteristic for that company given its combination of characteristics.
- The length of the vector that represents each variable indicates the discriminating power of the variable in separating the companies. Shorter vectors are those with the greatest discriminatory power (as long as their information is adequately represented on the plot).

PARALLEL COORDINATES

This method, proposed by Alfred Inselberg (1992), is a visualization system for presenting the n-dimensions in a 2-dimensional system. Each vertical axis represents an attribute that can be continuous or categorical. In addition, each of these axes can have its scale or be defined on a single scale. If the analysis has its own scale on each of the axes, a visualisation of hyper-surfaces can be made and the performance of the data can be analysed as a whole. When the analysis has a common scale, the relationship between the variables can be analysed. Both integrated techniques (HJ-Biplot and Parallel Coordinates) maximize success in the interpretation of the results (Galindo, Vicente, & Araya, 2009). Also, it is a comparison with the Cartesian axes, since instead of having one vertical and one horizontal, you have each of the variables in a vertical one in a parallel way and it is faster to understand. This method of representation is easy to do and does not require advanced knowledge of mathematics.

Added Value of Using HJ-Biplot, External Logistic Biplot and Parallel Coordinates

The added value of the HJ-Biplot, External Logistic Biplot and Parallel Coordinates to the question How to identifying sustainability gradients? is the statistical multivariate analysis. The combination of HJ-Biplot and Parallel Coordinates for the low dimensional representation is useful for better visualization of the results and for a better understanding of sustainability behaviour. The External Logistic Biplot is useful to identify and explore the similarity among individuals (companies in our case), to describe variables responsible for the similarities and to summarize the information to facilitate the interpretation of the data. The Biplots methods permit the joint plotting, in a reduced dimension of the rows where we can observe the explained variance of the analysis performed, the correlation between variables and individuals with similar sustainability characteristics.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

METHODOLOGY

Data

The top 500 companies with the highest revenues worldwide and published annually by FORTUNE magazine were considered for the present study. Concretly, the companies from the 2016 list were selected, and these companies' sustainability reports were subsequently searched in the database on the GRI website for the same year using version G4 as a guide. The indicators were obtained through a review of the companies' sustainability reports and websites. All of the FORTUNE 500 companies were searched, but only 159 of them reported using the G4 guide. The process of obtaining the indicators and the development of the database was carried out between April and May 2020. Subsequent updates were not taken into account for this research.

A total of 159 sustainability reports were downloaded, belonging to companies that are on the Global FORTUNE 500 list and have made their report with the GRI G4 guide. One of the reasons why the companies on this list were chosen is because of the variety of sectors, types of companies and because these types of companies tend to be more profitable and more exposed to public opinion. Also, they are under greater institutional pressure, because they are more visible (Garcia-Sanchez, Cuadrado-Ballesteros, & Frias-Aceituno, 2016).

Each company is caracterized by the following variables:

- Country: country in which the organization's central operating offices are located.
- Region: region classification according to the country of operation, in which sector it is located. The classification options in the GRI database are Africa, Asia, Europe, Latin America and the Caribbean, North America and Oceania.
- ISIC sector: classification of the business sector provided by the International Standard Industrial Classification (ISIC) of all economic activities of the International Labour Organization.
- Type of company: type of operation of the company according with the classification used by the GRI database. Three categories for classifying their reports given by the GRI comprises were herein considered: Private company, State-Owned Company and Subsidiary.
- Size: Dimension of the company according to GRI in the following way: Small and Medium Enterprises (less than 250 employees), Multinational Enterprises (equal or more than 250 employees and which are also Multinational Companies) and Large Enterprise (more than 250 employees).
- Level of application: indicator related to the Sustainability Reporting Framework and applied to a report. In the case of the G4 version, there are three levels: In accordance – Core, In accordance – Comprehensive, and Undeclared.
- GRI G4 indicators: 91 specific indicators in three categories, Economic, Environmental and Social . Predictor variables to be used in the ELB. See Appendix 1.

The data used in this chapter were organised in an **I**x**J** binary data matrix in which the **I** rows correspond to the companies and the **J** columns correspond to GRI G4 indicators.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Data Analysis

Firstly, a characterization of the companies used in the present study is summarily presented using various type of graphs such as, bar chart, pie chart and radar chart. In this part we used the following variables: country, region, ISIC sector, type of company, size and level of application.

Then, the predictor variables (GRI G4 indicators) were used for the HJ-Biplot and parallel coordinates analysis, to do that, a transformation was carried out on the data, thus obtaining the averages of the presences of the indicators concerning the different nominal variables. In this way, new **I**x**J** matrices were obtained where the individuals now corresponded to the different categories of the nominal variables and the columns the frequency averages in each of the indicators.

The ELB and HJ-Biplot were constructed using MULTBIPLOT: A package for Multivariate Analysis using Biplots (Vicente-Villardón, 2016). This package can be downloaded at http://biplot.dep.usal.es/multbiplot/download/

RESULTS

Caracterization of the Collected top Companies

Figure 1 shows the countries with the highest percentage of companies, with the United States standing out with 21% of the total. It is followed by the Asian countries of China and Japan with 12% and 11% respectively. Other countries on the Asian continent also stand out, such as the Republic of Korea with 6% of the total. Among the European countries, France, Germany and Sweden stand out with 6%, 8% and 4% respectively.

Taking into account the region to which the companies belong, most of them are companies from the Asian continent, with 36% of the total. Companies from Europe account for 35%, companies from the Americas account for 25% and companies from Oceania account for the smaller percentage (3%). See Figure 2.

Concerning on the application level of the indicators in the companies' reports, 51% of the companies are In accordance-Core, another 24% are Undeclared, 21% In accordance-Comprehensive and the remaining 4% have none (Figure 3).

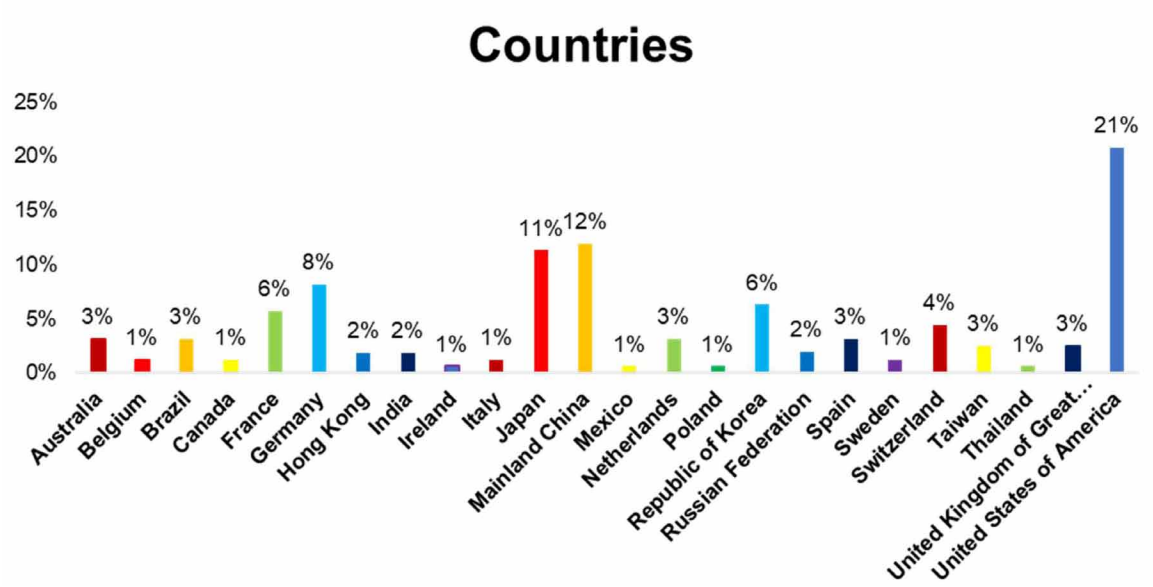
Figure 4 shows that companies with financial and insurance activities top the list of the most frequent in this study, with 23% of the total. This is followed by mining activities, with 19%, then commercial activities companies with 16%, followed by companies with information and communication activities with 15%. Manufacturing companies also account for 14% of the total, while other service activities account for 7%. Transport and storage companies, health activities and social services and construction companies have lower percentages, with 4% and 1% respectively.

Private sector companies account for 81% of the total. Only 16% correspond to public sector (State-Owned) companies and the remaining 3% to Subsidiary companies. See Figure 5.

This study found that 51% of the companies in the sample LARGE, while 48% Multinational enterprise (MNE). Only 1%, that is, one single company is Mid-size enterprises (SME). See Figure 6.

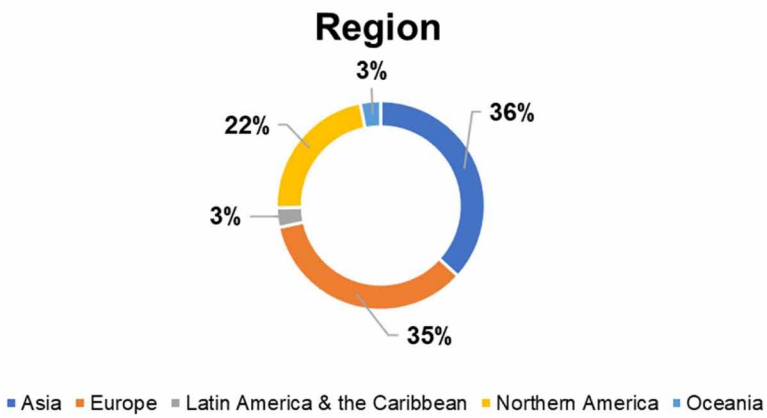
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Companies distribution by country



In terms of economic indicators, EC1 (Economic generated and distributed) is one of the most frequently reported in four regions (91.4% in Asia and North America), except in Latin America, which most frequently reports EC2 (Finantial implications due to the climate change). Another indicator that stands out for having a high frequency in all regions is the EC7 (Report the extent of development of significant infrastructure investments and services supported), the highest value in frequency is held by companies in the Oceania region, as it is reported by 80% of companies. A similar pattern can also be observed in Asia, Europe and North America in terms of the similarity of frequencies of the indicators reported. See Figure 7.

Figure 2. Companies distribution by region



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Companies distribution by application level

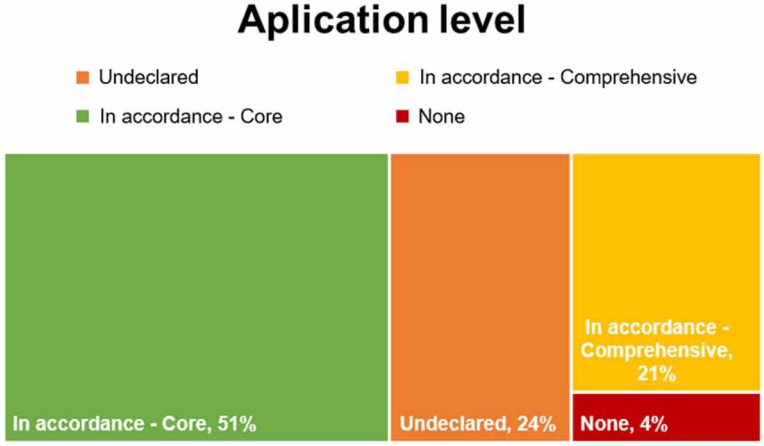
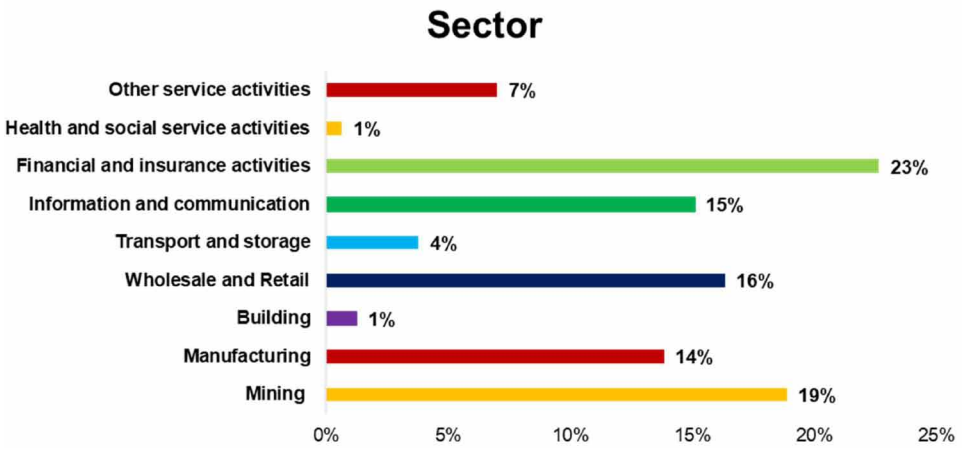


Figure 8 shows the indicators of the environmental dimension, where the indicator EN3 (Energy consumption within the organization) is one of the most reported, as well as EN8 (Total water withdrawal by source), EN15 (Direct Greenhouse Gas Emissions), and EN23 (Total weight of waste by type and disposal method). Once again, Europe and North America show similar behaviour.

Figure 9 shows labour practices indicators, where LA1(Total number and rate of employee turnover by age group, gender, and region), LA6 (Injuries, absentee rate and fatalities by region and by gender), and LA 12 (Composition of governance bodies and breakdown of employees per category according to gender, age group, minority group membership, and other indicators of diversity) are the most widely reported in most regions.

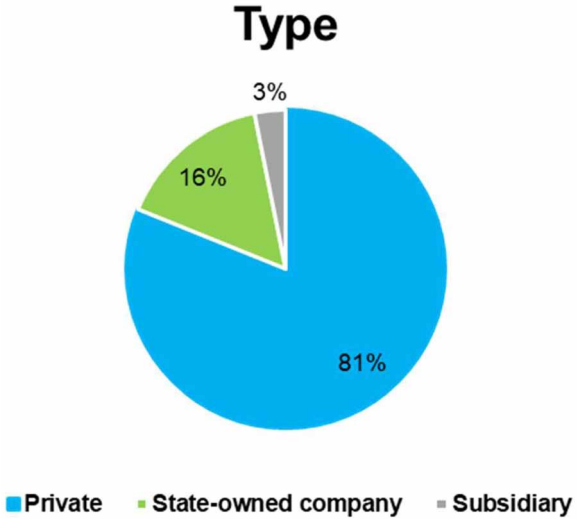
Concerning human rights indicators, differences can be seen between regions. It is clear that there is no pattern and that each region has different percentages of reporting indicators. In this dimension, the indicators HR2 (Hours of employee training in human right aspects), HR6 (Operations and suppliers identified as having significant risk for incidents of forced or compulsory labour, and measures

Figure 4. Companies distribution by ISIC sector



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Companies distribution by type

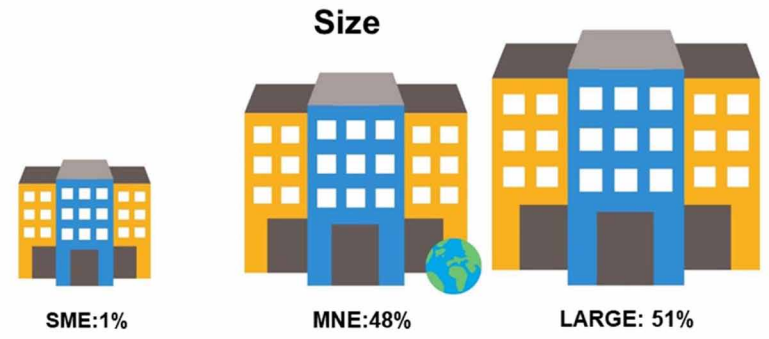


to contribute to the elimination of forced or compulsory labour), and HR10 (Percentage of significant suppliers and contractors that have undergone screening on human rights and actions taken) stand out, as most of them present high reporting percentages. See Figure 10.

In the Society dimension, none of the regions behaves similarly to each other, but as on previous occasions, some indicators stand out in all the regions, such as SO1 (Report the percentage of operations with implemented local community engagement, impact assessments, and development programs), SO4 (Percentage of employees trained in organization's anti-corruption policies and procedures), and some others, such as Latin America and Oceania, SO8 (Monetary value of significant fines and total number of non-monetary sanctions for noncompliance with laws and regulations), and SO10 (Negative social impacts in the supply chain and actions taken) stand out. See Figure 11.

Figure 12 shows a pattern in the reporting of certain Product Responsibility indicators, such as PR1 (Assessment of the health and safety impacts of product and service categories), PR3 (Requirements for product and service information and labeling), PR5 (Approach to stakeholder engagement, key topics and concerns raised

Figure 6. Companies distribution by size



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Economic indicators by region

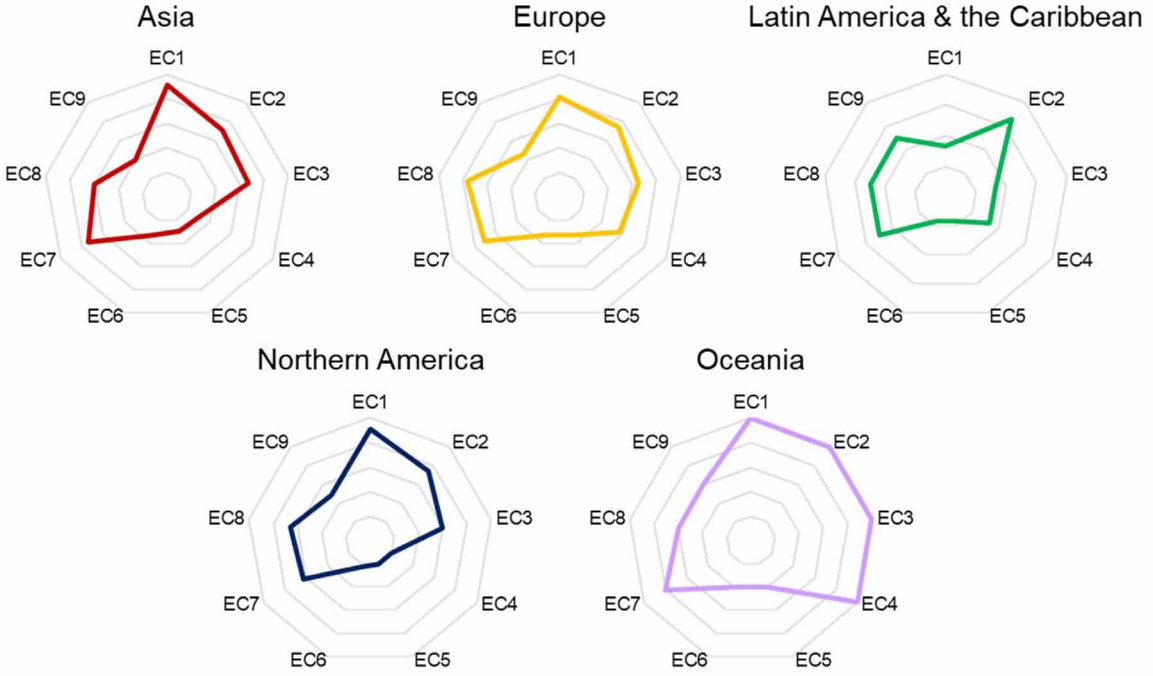
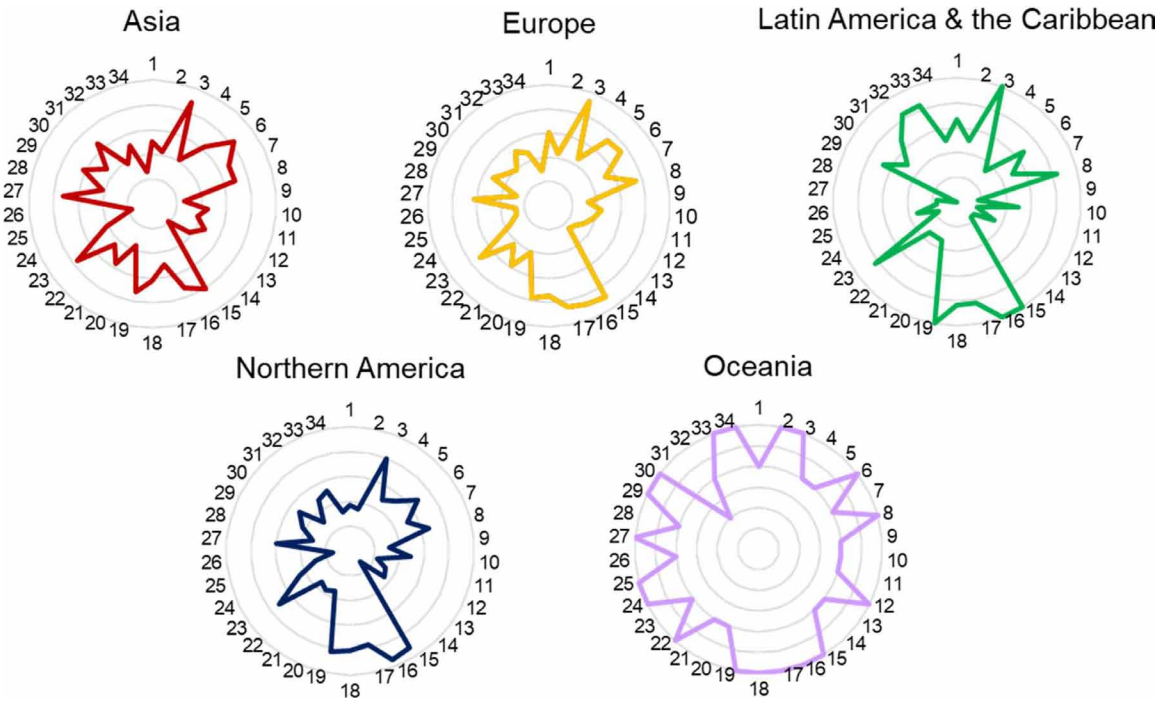


Figure 8. Environmental indicators by region



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Labour Practices indicators by region

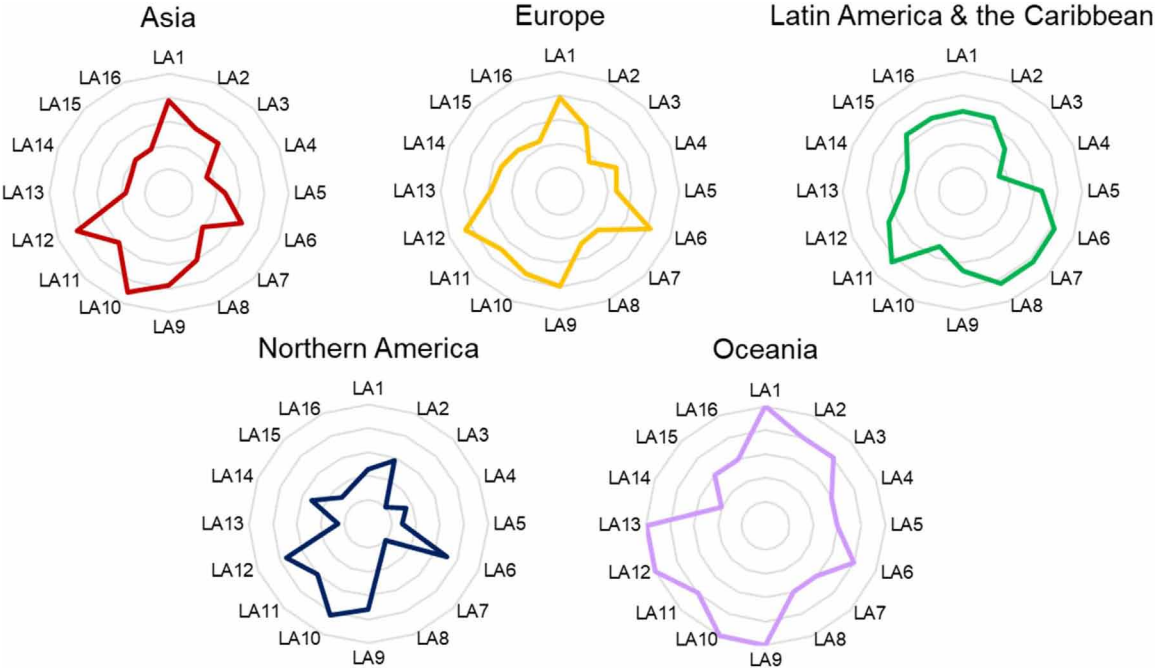
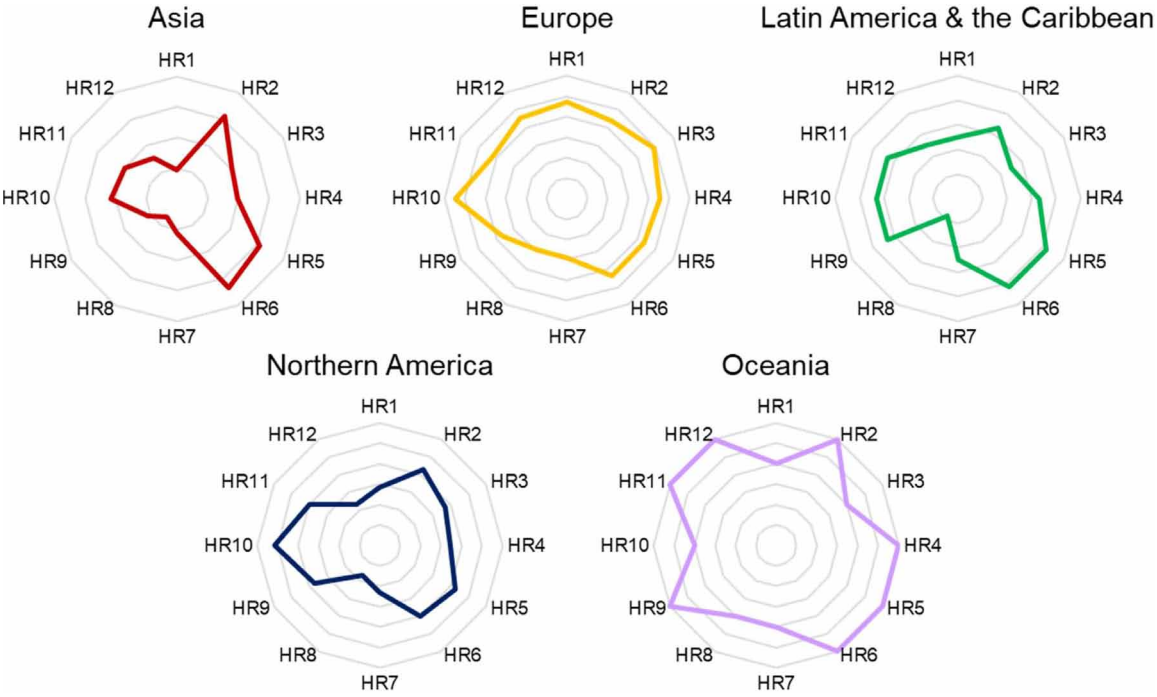
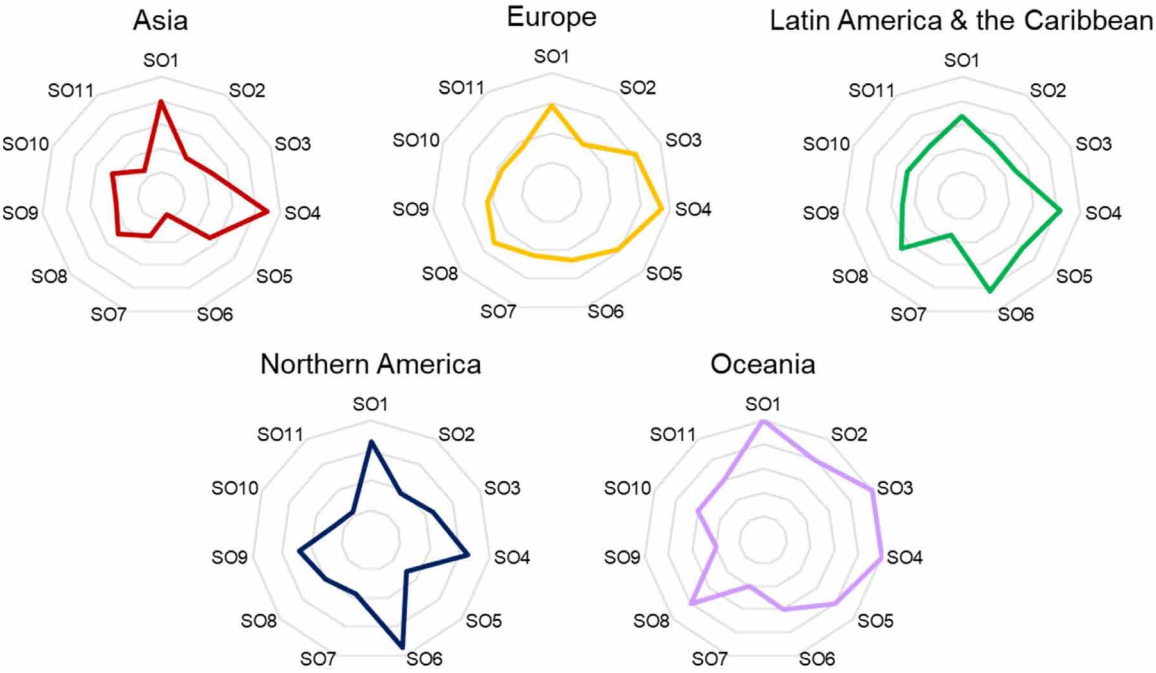


Figure 10. Human Rights indicators by region



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Society indicators by region



Key topics and concerns raised), and PR8 (Substantiated complaints concerning breaches of customer privacy and losses of customer data) with the highest reporting percentages in the regions. PR6 (Activities, brands, products, and services), and PR2 (Incidents of non-compliance concerning the health and safety impacts of products and services) indicators are the least reported in the regions.

Sustainability Gradients

As a means of obtaining the main gradients of sustainability and their relation to the observed characteristics and graphical representation, an ELB was constructed. The first principal plane (2D solutions) accounts for 19.70% of variability, and the first eigenvalue is 5.62. See Table 1.

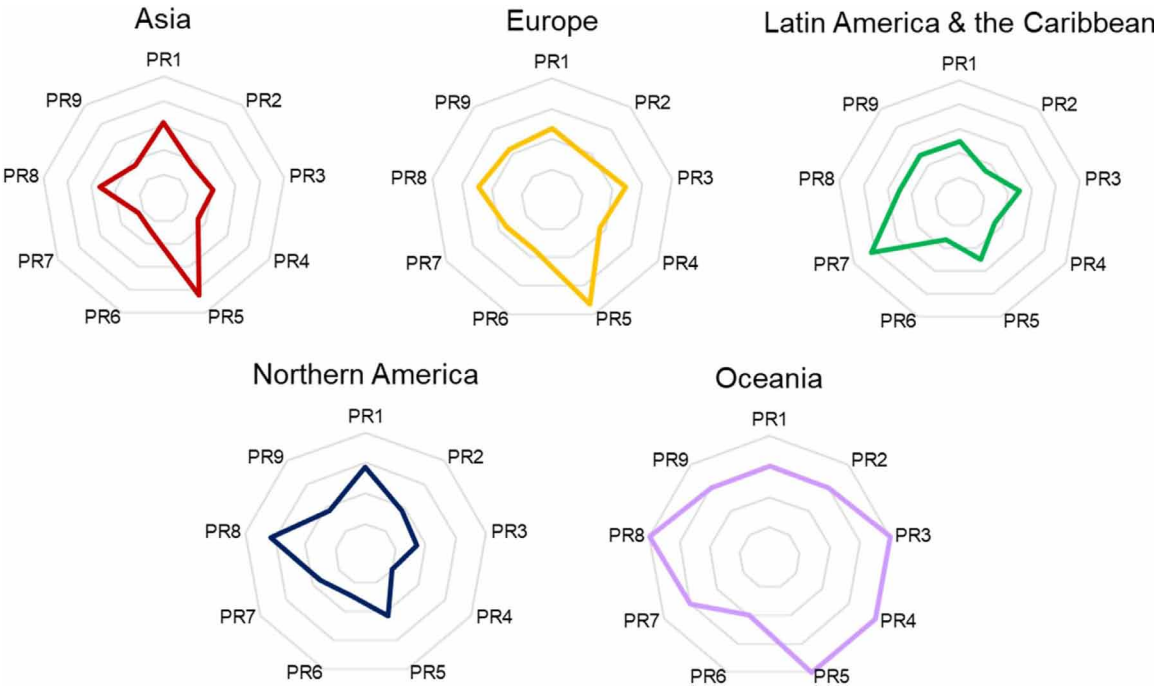
The overall goodness of fit (average of the goodness of fit for each variable) as a percentage of correct classification in the biplot is 77.28%, and the percentage of variables correctly classified, according to the results, is over 70% in the majority of cases, so there is a good prediction of the absence or presence of each GRI indicator, so the two-dimensional solution is sufficient to explore the main features of the data. See Table 2.

As for the sustainability gradient interpretation, the cosines of the angles are useful to interpret how variables are related to each gradient. In summary, there are two main gradients. See Table 3.

This first gradient is associated with proportion of senior management hired from the local community, water sources significantly affected by withdrawal of water, IUCN red list species and national conservation list species with habitats in areas affected by operations, emissions of ozone-depleting substances (ODS), water discharge by quality and destination, transport of hazardous waste, water bodies affected by water discharges and/or runoff, reclaimed products and their packaging materials, minimum notice

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Product Responsibility by region



periods regarding operational changes, ratio of basic salary and remuneration of women to men, negative social impacts in the supply chain and actions taken, and management approach and its components. From the ELB, Figure 13, the companies located on the right side of the factorial axis 1 report more of these indicators related to the first gradient than those located on the left side. The second sustainability gradient is associated with the environmental dimension related to energy consumption within the organization, GHG emissions, and reduction of GHG emissions. From Figure 13, the companies located bottom the factorial axis 1 have greater values for these indicators as opposed to those located above of factorial axis 1.

To create clusters of companies with similar profiles of sustainability, the Ward's method of hierarchical clustering with the ordination scores from the ELB was used. In Figure 14, three clusters are highlighted. Cluster 1 comprises companies with fewer indicators, therefore low sustainability. This cluster comprises 25% of the companies in total and also, 40% of the companies in this cluster are from the financial services sector. Another characteristic to highlight is that the 37.5% are Asian companies, of the private type and large size. However, 32.5% are also North American companies. Cluster 2 comprises

Table 1. Logistic biplot with external information

Retained axes and explained variance			
Axes	**Eigenvalue**	**% of variance**	**Cumulative %**
1	5.6174	14.6607	14.6607
2	1.9294	5.0354	19.6961

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Percentage of variables correctly classified

Variable	Standard deviation	p-value	R2	% correct classification
EC6	77.14	0	0.53	81.76
EN3	56.25	0	0.53	91.20
EN9	160.04	0	0.80	87.42
EN14	85.71	0	0.60	88.05
EN15	105.99	0	0.78	94.34
EN16	97.26	0	0.71	90.57
EN17	77.70	0	0.54	83.02
EN18	91.44	0	0.62	85.54
EN19	84.28	0	0.61	88.05
EN20	90.13	0	0.58	78.62
EN22	90.85	0	0.58	74.84
EN25	85.84	0	0.58	84.91
EN26	106.38	0	0.67	87.42
EN28	82.05	0	0.55	80.50
LA4	75.77	0	0.51	77.99
LA13	84.41	0	0.55	84.91
LA15	81.09	0	0.53	74.84
LA16	86.84	0	0.56	81.13
HR12	103.06	0	0.64	78.62
SO11	106.60	0	0.65	84.91

23.27% of the total number of companies, 24% of which are companies in the mining sector. Another sector which stands out in this cluster is the communication and information sector. This cluster is also comprised 51.4% of Asian companies and fewer companies from the Latin American region. An interesting fact is that 54.10% of these companies are large. Cluster 3 is the largest one and comprises 51.5% of the total number of companies. In this cluster there are characteristics such as companies from the wholesale and retail trade sector; the highest number of companies (39%) are from the European region and of the private sector (92%). In this cluster, 54% of the companies are MNE and 45.10% are large. Therefore, this cluster seems to comprise the most sustainable companies in this study.

Regional Analysis

The distribuition of the categories reported by the top companies turns out to be different around the world. Indeed, following Figure 15 where the six vertical axis are representing the percentage of companies of a given region (AS, EU, LC and NA) that reported indicators into the economic (EC), environment (EN) and social (LA, HR, SO and PR) categories respectively, the four regions exhibits different profiles. This graph uses a single scale, i.e. the highest percentage of indicators reported by companies is 61% in labour practices and is from Latin America, and the lowest is 34% in human rights from North America. The purpose of this graph is to visualise in which indicator each of the regions studied stands out.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Cosines of the angles

Variable	Cos Axis 1	Cos. Axis 2	Gradient
EC6	0.859	0.512	1
EN3	-0.339	0.941	2
EN9	0.788	-0.616	1
EN14	0.856	-0.516	1
EN15	-0.151	0.989	2
EN16	-0.206	0.979	2
EN17	-0.289	0.957	2
EN18	-0.368	0.930	2
EN19	-0.270	0.963	2
EN20	0.833	-0.553	1
EN22	0.959	-0.284	1
EN25	0.908	-0.419	1
EN26	0.995	-0.097	1
EN28	0.747	-0.664	1
LA4	0.810	0.587	1
LA13	0.910	0.414	1
LA15	0.973	-0.230	1
LA16	0.861	0.509	1
HR12	0.927	0.375	1
SO11	0.752	0.659	1

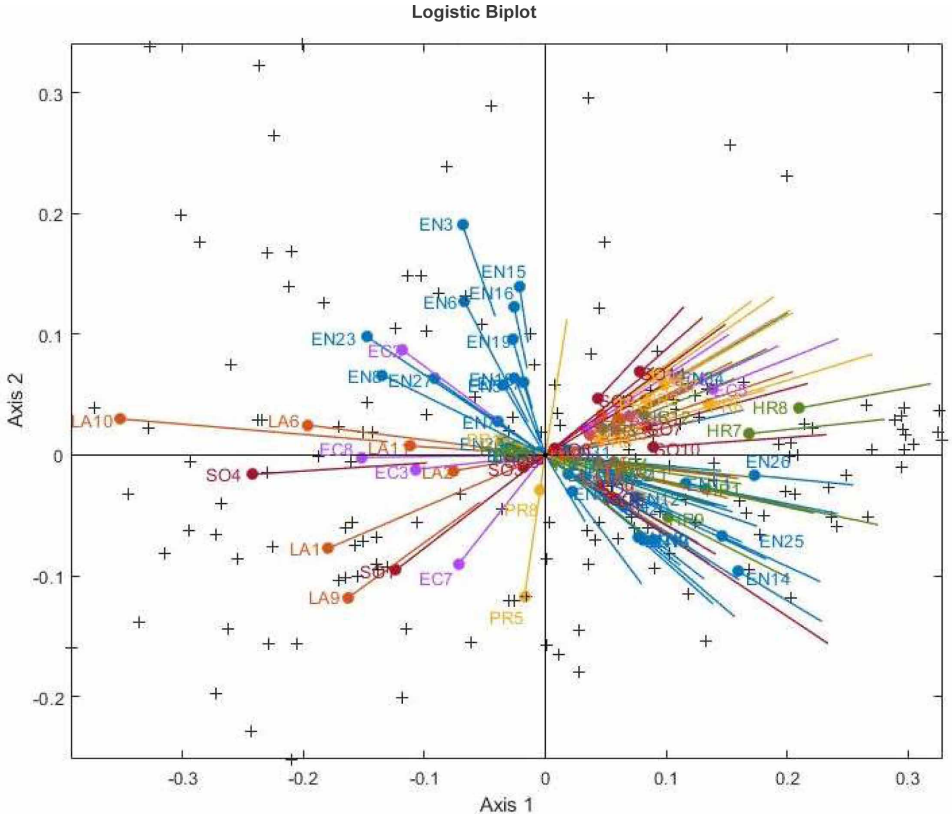
Looking at the position on the vertical axes related to the human rights (HR), Asia stands out from the rest by a large margin. It is followed by Latin America, Europe and finally North America. In the category of society (SO), Latin America stands out and distance from the second place occupied by the Europe region, followed by North America and Asia in the last places opposed to the previous category. In the economic field (EC), however, Europe stands out from the other regions. North America is the region with the lowest performances in terms of having companies reporting indicators in most categories, except for both the economic category and the society. Moreover, in the parallel coordinates in Figure 15, given the proximity of the points on the vertical axes related to the category of environment (EN) and the product responsibility (PR), there is not much difference between the four regions in terms of performance in reporting the indicators in these two categories.

The following results were obtained by analysing the nominal region variable and the different aspects in which the indicators are classified. It is possible to capture an explained variance in the first two axes of 87%. See Table 4.

In the results of the HJ-Biplot, each of the regions can be seen clearly separated and also results that coincide with the previous analysis. The aspects with the greatest variability are PP (public policy) and PS (products and services). Therefore, an HJ-Biplot complements the results of the previous parallel coordinate analysis, where Asia is characterised by higher averages in the reporting of indicators in human rights aspects, but the specific indicators that characterise it are TF (Forced Labour), TI (Child

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. External logistic biplot, based on sustainability data from Global FORTUNE 500 companies



Labour), EDP (Supplier Assessment), EV (Overall Assessment), MDS (Security Measures) and LDA (Freedom of Association and Collective Bargaining). It can also be noted that Latin America stands out in aspects of labour practices and society, in the specific aspects SYS (Health and Safety at Work), MDR (Mechanisms for Complaints about Labour Practices), CR (Regulatory Compliance) and MRS (Mechanisms for Complaints about Social Impact). While North America is characterised by environmental aspects (AG: Water, BIO: Biodiversity and TRA: Transport). Europe is not well represented on this level, but the aspects that best characterise it are seen to include environmental categories (EN: Energy and ER: Effluents and Waste), product responsibility (SSC: Customer Health and Safety and PC: Customer Privacy) and economic aspects (PM: Market Presence, DE: Economic Performance and CE: Economic Consequences). See Figure 16.

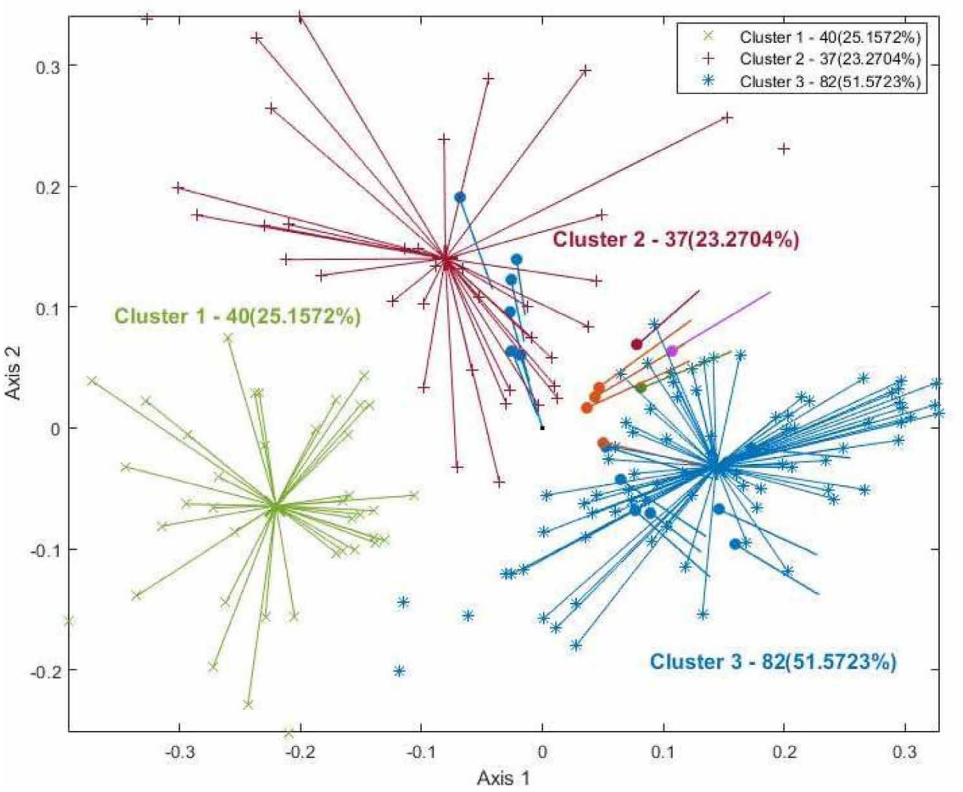
The same analysis could also be made for countries, economic sector, type of company, and company size, using the same statistical procedures but for reasons of volume they are not presented in this chapter.

DISCUSSION

The largest number of companies on the studied list are North American, European and Asian. These regions are home to more large and multinational companies, and so they are on the list of companies with the highest revenues and tend to report their sustainability practices more frequently. While there

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. External logistic biplot with clusters



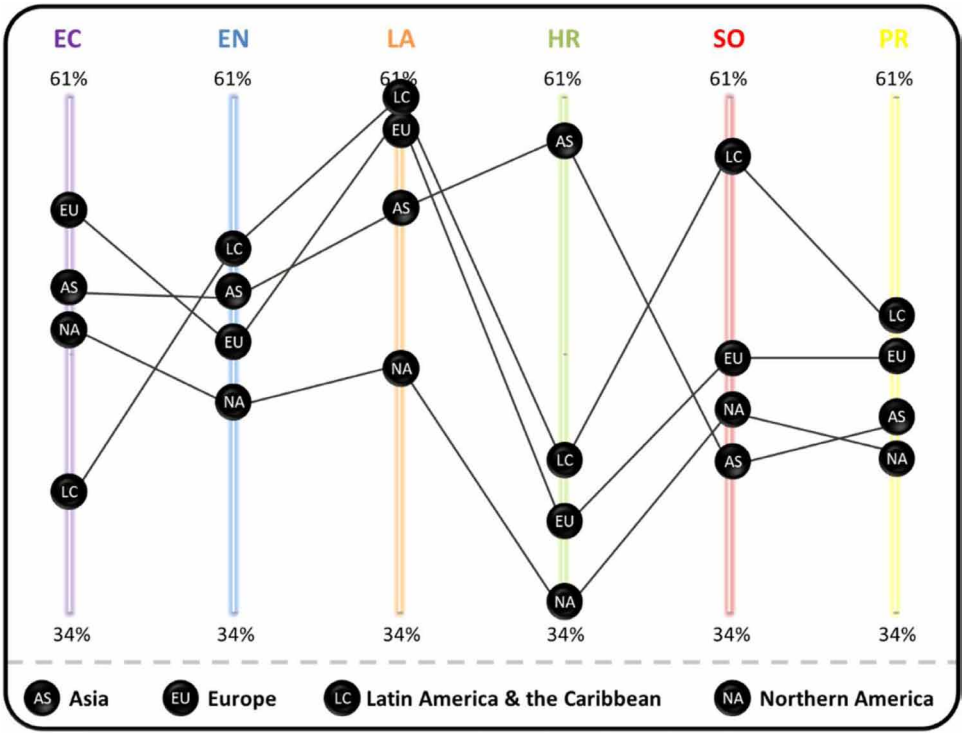
are fewer in Latin America and Oceania, these sectors have smaller numbers of large and with high-revenue companies and are therefore not listed as the highest income earners. The predominant sector is financial services and insurance, which may indicate that they are the most concerned with reporting their sustainability practices over that period, even if this does not mean that they are the most sustainable. The vast majority of companies are private, which indicates that these types of companies may be the ones that place the greatest emphasis on reporting sustainability practices and being transparent to stakeholders. Unlike the study carried out in 2015 where the cluster was defined as the most sustainable, in which the majority of large companies belong to the water and energy sector (Tejedor-Flores, 2015).

The most commonly reported indicators or those that companies consider relevant to present in their reports are energy consumption within the organization (EN3), and direct GHG emissions (EN15). Those companies with the least reporting are IUCN red list species and national conservation list species with habitats in areas affected by operations (EN14) and incidents of violations involving rights of indigenous peoples (HR8). This fact may be because, as there are more companies in the financial sector, their operations do not directly impact on the habitat of endemic species and similarly, the rights of indigenous people are not affected during operations.

Similar studies can be found where the External Logistic Biplot was also applied with success. For instances, in Tejedor- Flores (2015) two main sustainability gradients were found. The first one is related to EN11, EN12, EN20, LA5, LA8, HR3, HR5, HR6 and HR7. and is associated with biodiversity, emission of greenhouse gases, the relationship of the company with their workers, human rights, the fighting

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Parallel Coordinates for descreving the percentage of companies of four regions that have been reporting indicators in each six categories



against child labour and forced labour. The second sustainability gradient comprises the variables: LA15, HR10, HR11, SO9 and SO10 and is associated with the social dimension related to employment, human rights evaluation and local communities. Tejedor- Flores also found the cluster with most sustainable companies (Cluster 3) during 2011, 2012 and 2013. These companies are then associated with a great concern for biodiversity conservation and air pollution prevention as well as a strong relationship with their employees, while maintaining respect for and promotion of human rights and the fighting against child labour and forced labour. Cluster 3 comprises 140 companies (37.14%) of which 90% belong to the Water and Energy Sector (36.43%). This is the group with the highest report qualifications and sustainability reports verified by the GRI (i.e. GRI-checked).

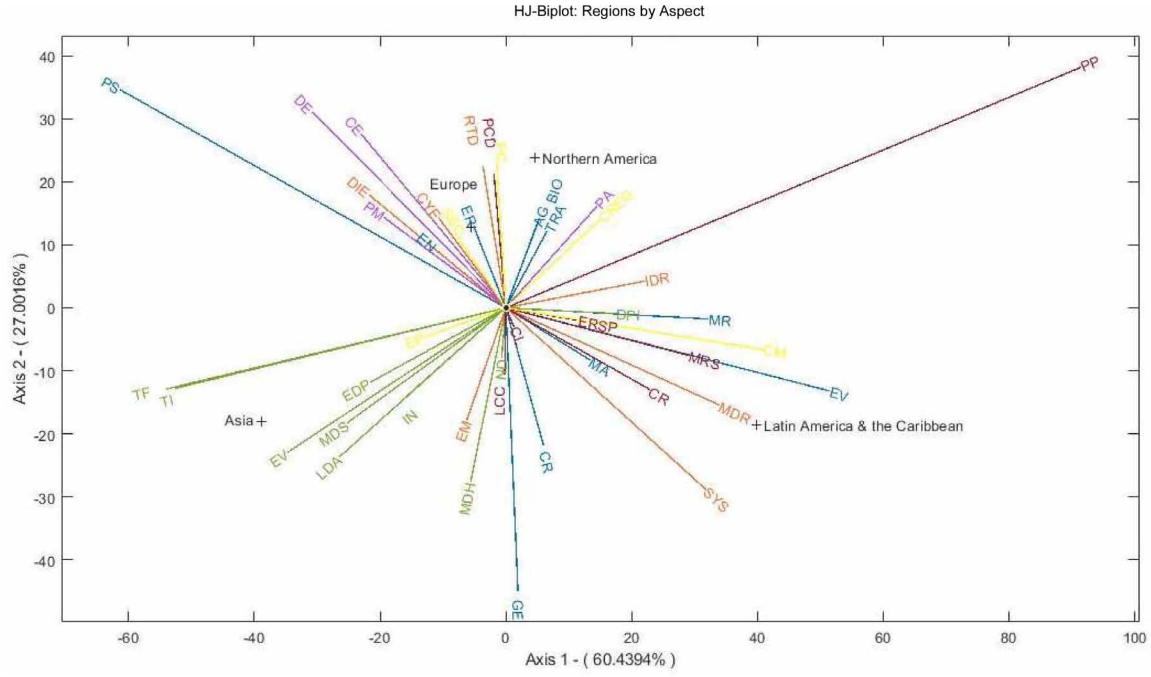
For the Latin American region, Urruticoechea & Vernaza (2018), based on similar methodology, observed that the indicators of labour practices and society are more reported, while for North America they are those of economic and environmental aspects. Therefore, the results obtained in the present

Table 4. Explained variance of regions by aspect

Retained axes and explained variance			
Axes	Eigenvalue	% of variance	Cumulative %
Axis 1	10438.88	60.439	60.439
Axis 2	4663.61	27.002	87.441

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. HJ-Biplot of regions by aspect



study are in accordance with those mentioned, as Latin America is strong in terms of labour practices and social indicators.

The European region shows that the strength in reporting is in economic aspects and the second with the most companies reporting on labour practice aspects. This is in accordance with the results by Vicente-Galindo et al. (2015), in which those authors found that the most reported indicators were firstly economic, environmental and thirdly the Labour Practices indicators. In that study, Spain stands out as one of the countries with the highest percentage of presence in its GRI reports of all the countries in the study.

Another aspect to mention is that according to the type of company there are different reporting habits, which up to now are unknown. Subsidiary companies report mostly on economic matters, while private companies are more inclined to report on Labour Practices, Society and Human Rights. Interestingly, it is the state-owned companies that are outstanding in the economic and product responsibility areas. Cubilla-Montilla (2019) mentions that the impact derived from this information practice generates benefits, mainly to improve its image, reputation and possible new interested parties in financing with better conditions.

CONCLUSIONS

According to the results obtained in the present study for the Fortune 500 companies (2016), the main conclusions are:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. Only 159 use the G4 guide to report their sustainability activities.
2. Two sustainability gradients are identified. The main gradient contains indicators from the economic, environmental, labour practices, society and human rights categories. While the second gradient is defined only by indicators of the environmental category. In these two gradients, the companies are ordered, so it is possible to define their sustainability reporting trend.
3. Three clusters of companies were found which could be called the most sustainable: wholesale and retail companies, European region and private type. Theleast sustainable comprises mostly financial services companies, from the Asian region and private type.
4. The most frequently reported indicators by companies include both the Environmental (EN3, EN15 and EN16) and Labour Practices (LA10 and LA12) among others. They are considered to be the most widely reported as they apply to practically any sector of activity. Those reported less frequently belong to the Environmental category (EN14, EN25 and EN26).Maybe they are less reported because they do not apply to any type of organisation.
5. Although the same results could be reached with other statistical methods, in this chapter three exploratory techniques are mainly highlighted. The combination of methods such as HJ-Biplot and parallel coordinates for the low dimensional representation was useful for better visualization of the results. The External Logistic Biplot was useful to identify and explore the similarity among objects (companies), to describe variables responsible for the similarities and to summarize the information in latent gradients to facilitate the interpretation of the data.

ACKNOWLEDGMENT

This work was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT – Fundação para a Ciência e a Tecnologia), reference UIDB/04106/2020.

REFERENCES

Artaraz, M. (2002). Teoría de las tres dimensiones de desarrollo sostenible. *Ecosistemas (Madrid)*, *2002*(2). http//www.aeet.org/ecosistemas/022/informe1.htm

Brundtland, G. H. (1987). *Report of the World Commission on Environment and Development: Our Common Future (‘The Brundtland Report’).* doi:10.9774/GLEAF.978-1-907643-44-6_12

Buck, B., Espinach, L., & Söderberg, S. (2014). *GRI G4 Guidelines and ISO 26000:2010 How to use the GRI G4 Guidelines and ISO 26000 in conjunction*. Retrieved from https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/iso-gri-26000_2014-01-28.pdf

Bybee, R. W. (1991). Planet Earth in Crisis: How Should Science Educators Respond? *The American Biology Teacher*, *53*(3), 146–153. doi:10.2307/4449248

Cubilla-Montilla, M., Rodriguez, C., & Ortega, E. (2015). Análisis de los indicadores de sostenibilidad del Global Reporting Initiative. Una mirada desde el Biplot Logístico. *Revista Centros*, *3*, 96–114.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Davis, K., & Blomstrom, R. (1980). *Business and society. Concepts and policy issues* (4th ed.). McGraw Hill.

Demey, J. R., Vicente-Villardon, J. L., Galindo-Villardon, M. P., & Zambrano, A. Y. (2008, December 15). (2208). Identifying molecular markers associated with classification of genotypes by external logistic biplots. *Bioinformatics (Oxford, England)*, *24*(24), 2832–2838. doi:10.1093/bioinformatics/btn552 PMID:18974073

Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*(3), 211–218. doi:10.1007/BF02288367

Gabriel, K. R. (1971). The biplot - graphic display of matrices with applications to principal component analysis. *Biometrika*, *58*(3), 453–467. doi:10.1093/biomet/58.3.453

Gabriel, K. R. (1995). Biplot display of multivariate categorical data, with comments on multiple correspondence analysis. In *Recent Advances in Descriptive Multivariate Analysis*. Oxford Science Publications.

Gabriel, K. R. (1998). Generalised bilinear regression. *Biometrika*, *85*(3), 689–700. doi:10.1093/biomet/85.3.689

Gabriel, K. R., & Zamir, S. (1979). And Zamir, S. (1979). Lower rank approximation of matrices by least squares with any choice of weights. *Technometrics*, *21*(4), 489–498. doi:10.1080/00401706.1979.10489819

Galindo, P. (1986). Una representación simultánea: HJ Biplot. *Qüestiió. Quaderns d'Estadística, Sistemes, Informàtica i Investigació Operativ*, *10*, 13–23.

Galindo, P., Vicente, P., & Araya, C. (2009). Biplot Versus Coodernadas Paralelas. In Memoria del XXIII Foro de Estadística (pp. 55-59). Academic Press.

Global Reporting Initiative. (2013). *Guide for the preparation of sustainability reports*. G4. Retrieved from https://www.globalreporting.org/resourcelibrary/Spanish-G4-Part-One.pdf

Gómez Gutiérrez, C. (2018). United Nations Educational, Scientific and Cultural Organization (UNESCO) [El desarrollo sostenible: conceptos básicos, alcance y criterios para su evaluación]. *Tema*, III.

Gower, J. C., & Hand, D. (1996). *Biplots*. Chapman & Hall.

Ho, L. C. J., & Taylor, M. E. (2007). An empirical analysis of triple bottom-line reporting and its determinants: Evidence from the United States and Japan. *Journal of International Financial Management & Accounting*, *18*(2), 123–150. doi:10.1111/j.1467-646X.2007.01010.x

Holme, R., & Watts, P. (2001). Making Good Business Sense. *Journal of Corporate Citizenship*, *2001*(2), 17–20. doi:10.9774/GLEAF.4700.2001.su.00005

Inselberg, A. (1992). *The plane R2 with coordinate parallel. Computer Science and Applied Mathematics Departments. Tel Aviv*. University.

Institute of Chartered Accountants in England and Wales (ICAEW). (2004). *Sustainability: The Role of Accountants.* Sustainable Business Initiative Report.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nieto, A. B., Galindo-Villardón, M., Leiva, V., & Vicente-Galindo, M. (2014). A Methodology for Biplots based on bootstrapping with R. *Revista Colombiana de Estadistica*, *37*(2Spe), 367–397. doi:10.15446/rce.v37n2spe.47944

Patten, D. M. (1991). Exposure, legitimacy, and social disclosure. *Journal of Accounting and Public Policy*, *10*(4), 297–308. doi:10.1016/0278-4254(91)90003-3

Rodriguez Rojas, C. I. (2015). *Sostenibilidad en las empresas*. Recuperado 6 de julio de 2020, de Escuela de Organización InduStrial website: https://www.eoi.es/blogs/carollirenerodriguez/2012/05/20/sostenibilidad-en-las-empresas/

Rouf, M. A. (2011). The Corporate Social Responsibility Disclosure: A Study of Listed Companies in Bangladesh. *Business and Economics Research Journal*, *2*(3), 19–32.

Santolaria, C. J. S., & González, J. P. G. (2013). Las memorias de sostenibilidad y su divulgación. *Contaduría Universidad de Antioquia*, (57), 107-118. Retrieved from https://revistas.udea.edu.co/index.php/cont/article/view/15580/13523

Tejedor, N. (2015). *Análisis Multivariante de la Sosteniblidad en Brasil para los años 2011, 2012 y 2013, a través del Global Reporting Initiative (GRI) (Trabajo fin de máster)*. Universidad de Salamanca.

Tejedor-Flores, N., Galindo-Villardón, P., & Vicente-Galindo, P. (2017). Sustainability multivariate analysis based on the global reporting initiative (GRI) framework, using as a case study: Brazil compared to Spain and Portugal. *International Journal of Sustainable Development and Planning*, *12*(4), 667–677. doi:10.2495/SDP-V12-N4-667-677

United Nations. (2019). *Sustainable Development Goals*. Retrieved from https://sdgs.un.org/goals

Urruticoechea, A., & Vernaza, E. (2018). *Una aplicación de la metodología biplot logísico: análisis de la sostenibilidad empresarial*. XI Semana Internacional de la Estadística y la Probabilidad. Facultad de Ciencias Físico Matemáticas Puebla.

Vicente Galindo, P., Vaz, E., & de Noronha, T. (2015). How Corporations Deal with Reporting Sustainability: Assessment Using the Multicriteria Logistic Biplot Approach. *Systems*, *3*(1), 6–26. doi:10.3390ystems3010006

Vicente Villardón, J.L. (2015). *Multbiplot: A package for Multivariate Analysis using Biplots. Departamento de Estadística.* Universidad de Salamanca. http://biplot.usal.es/ClassicalBiplot/index.html

Vicente-Villardon, J. L., Galindo-Villardon, M. P., & Blazquez-Zaballos, A. (2006). *Logistic Biplots in Multiple Correspondence Analysis and Related Methods*. Statistics in Social and Behavioural Sciences Series.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Corporate Social Responsibility: Management concept whereby companies integrate economic, social and environmental concerns in their business operations and interactions with their stakeholders.

Global Fortune 500 Companies: Fortune magazine's list of the 500 largest companies in the world.

HJ-Biplot: A multivariate graphical representation of row and column markers, chosen in such a way that they can be overlapped in the same reference system with maximum quality of representation.

Logistics Biplot: A linear biplot for binary data in which response along the dimensions is logistic.

Multivariate Analysis: Set of statistical techniques used for analysis of data that contain more than one variable.

Parallel Coordinates: A visualization system for presenting the n-dimensions in a 2-dimensional system. Each vertical axis represents an attribute that can be continuous or categorical.

Sustainability Report: Document that present a company's environmental, social, and economic performance.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX 1

Table 4. GRI Indicator Table

G4 Disclosure	GRI Standard Number	Disclosure Number	Disclosure Title
G4-EC1	GRI 201	201-1	Direct economic value generated and distributed
G4-EC2	GRI 201	201-2	Financial implications and other risks and opportunities due to climate change
G4-EC3	GRI 201	201-3	Defined benefit plan obligations and other retirement plans
G4-EC4	GRI 201	201-4	Financial assistance received from government
G4-EC5	GRI 202	202-1	Ratios of standard entry level wage by gender compared to local minimum wage
G4-EC6	GRI 202	202-2	Proportion of senior management hired from the local community
G4-EC7	GRI 203	203-1	Infrastructure investments and services supported
G4-EC8	GRI 203	203-2	Significant indirect economic impacts
G4-EC9	GRI 204	204-1	Proportion of spending on local suppliers
G4-EN1	GRI 301	301-1	Materials used by weight or volume
G4-EN2	GRI 301	301-2	Recycled input materials used
G4-EN3	GRI 302	302-1	Energy consumption within the organization
G4-EN4	GRI 302	302-2	Energy consumption outside of the organization
G4-EN5	GRI 302	302-3	Energy intensity
G4-EN6	GRI 302	302-4	Reduction of energy consumption
G4-EN7	GRI 302	302-5	Reductions in energy requirements of products and services
G4-EN8	GRI 303	303-1	Water withdrawal by source
G4-EN9	GRI 303	303-2	Water sources significantly affected by withdrawal of water
G4-EN10	GRI 303	303-3	Water recycled and reused
G4-EN11	GRI 304	304-1	Operational sites owned, leased, managed in, or adjacent to, protected areas and areas of high biodiversity value outside protected areas
G4-EN12	GRI 304	304-2	Significant impacts of activities, products, and services on biodiversity
G4-EN13	GRI 304	304-3	Habitats protected or restored
G4-EN14	GRI 304	304-4	IUCN Red List species and national conservation list species with habitats in areas affected by operations
G4-EN15	GRI 305	305-1	Direct (Scope 1) GHG emissions
G4-EN16	GRI 305	305-2	Energy indirect (Scope 2) GHG emissions
G4-EN17	GRI 305	305-3	Other indirect (Scope 3) GHG emissions
G4-EN18	GRI 305	305-4	GHG emissions intensity
G4-EN19	GRI 305	305-5	Reduction of GHG emissions
G4-EN20	GRI 305	305-6	Emissions of ozone-depleting substances (ODS)
G4-EN21	GRI 305	305-7	Nitrogen oxides (NOX), sulfur oxides (SOX), and other significant air emissions
G4-EN22	GRI 306	306-1	Water discharge by quality and destination
G4-EN23	GRI 306	306-2	Waste by type and disposal method
G4-EN24	GRI 306	306-3	Significant spills
G4-EN25	GRI 306	306-4	Transport of hazardous waste
G4-EN26	GRI 306	306-5	Water bodies affected by water discharges and/or runoff
G4-EN27	NA	NA	NA
G4-EN28	GRI 301	301-3	Reclaimed products and their packaging materials
G4-EN29	GRI 307	307-1	Non-compliance with environmental laws and regulations
G4-EN30	NA	NA	NA
G4-EN31	Several	NA	NA
G4-EN32	GRI 308	308-1	New suppliers that were screened using environmental criteria
G4-EN33	GRI 308	308-2	Negative environmental impacts in the supply chain and actions taken
G4-EN34	GRI 103	103-2	The management approach and its components
G4-LA1	GRI 401	401-1	New employee hires and employee turnover
G4-LA2	GRI 401	401-2	Benefits provided to full-time employees that are not provided to temporary or part-time employees
G4-LA3	GRI 401	401-3	Parental leave

continued on following page

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4 Continued

G4 Disclosure	GRI Standard Number	Disclosure Number	Disclosure Title
G4-LA4	GRI 402	402-1	Minimum notice periods regarding operational changes
G4-LA5	GRI 403	403-1	Workers representation in formal joint management–worker health and safety committees
G4-LA6	GRI 403	403-2	Types of injury and rates of injury, occupational diseases, lost days, and absenteeism, and number of work-related fatalities
G4-LA7	GRI 403	403-3	Workers with high incidence or high risk of diseases related to their occupation
G4-LA8	GRI 403	403-4	Health and safety topics covered in formal agreements with trade unions
G4-LA9	GRI 404	404-1	Average hours of training per year per employee
G4-LA10	GRI 404	404-2	Programs for upgrading employee skills and transition assistance programs
G4-LA11	GRI 404	404-3	Percentage of employees receiving regular performance and career development reviews
G4-LA12	GRI 405	405-1	Diversity of governance bodies and employees
G4-LA13	GRI 405	405-2	Ratio of basic salary and remuneration of women to men
G4-LA14	GRI 414	414-1	New suppliers that were screened using social criteria
G4-LA15	GRI 414	414-2	Negative social impacts in the supply chain and actions taken
G4-LA16	GRI 103	103-2	The management approach and its components
G4-HR1	GRI 412	412-3	Significant investment agreements and contracts that include human rights clauses or that underwent human rights screening
G4-HR2	GRI 412	412-2	Employee training on human rights policies or procedures
G4-HR3	GRI 406	406-1	Incidents of discrimination and corrective actions taken
G4-HR4	GRI 407	407-1	Operations and suppliers in which the right to freedom of association and collective bargaining may be at risk
G4-HR5	GRI 408	408-1	Operations and suppliers at significant risk for incidents of child labor
G4-HR6	GRI 409	409-1	Operations and suppliers at significant risk for incidents of forced or compulsory labor
G4-HR7	GRI 410	410-1	Security personnel trained in human rights policies or procedures
G4 Disclosure	**GRI Standard Number**	**Disclosure Number**	**Disclosure Title**
G4-HR8	GRI 411	411-1	Incidents of violations involving rights of indigenous peoples
G4-HR9	GRI 412	412-1	Operations that have been subject to human rights reviews or impact assessments
G4-HR10	GRI 414	414-1	New suppliers that were screened using social criteria
G4-HR11	GRI 414	414-2	Negative social impacts in the supply chain and actions taken
G4-HR12	GRI 103	103-2	The management approach and its components
G4-SO1	GRI 413	413-1	Operations with local community engagement, impact assessments, and development programs
G4-SO2	GRI 413	413-2	Operations with significant actual and potential negative impacts on local communities
G4-SO3	GRI 205	205-1	Operations assessed for risks related to corruption
G4-SO4	GRI 205	205-2	Communication and training about anti-corruption policies and procedures
G4-SO5	GRI 205	205-3	Confirmed incidents of corruption and actions taken
G4-SO6	GRI 415	415-1	Political contributions
G4-SO7	GRI 206	206-1	Legal actions for anti-competitive behavior, anti-trust, and monopoly practices
G4-SO8	GRI 419	419-1	Non-compliance with laws and regulations in the social and economic area
G4-SO9	GRI 414	414-1	New suppliers that were screened using social criteria
G4-SO10	GRI 414	414-2	Negative social impacts in the supply chain and actions taken
G4-SO11	GRI 103	103-2	The management approach and its components
G4-PR1	GRI 416	416-1	Assessment of the health and safety impacts of product and service categories
G4-PR2	GRI 416	416-2	Incidents of non-compliance concerning the health and safety impacts of products and services
G4-PR3	GRI 417	417-1	Requirements for product and service information and labeling
G4-PR4	GRI 417	417-2	Incidents of non-compliance concerning product and service information and labeling

continued on following page

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4 Continued

G4 Disclosure	GRI Standard Number	Disclosure Number	Disclosure Title
G4-PR5	GRI 102	102-43	Approach to stakeholder engagement
		102-44	Key topics and concerns raised
G4-PR6	GRI 102	102-2	Activities, brands, products, and services
G4-PR7	GRI 417	417-3	Incidents of non-compliance concerning marketing communications
G4-PR8	GRI 418	418-1	Substantiated complaints concerning breaches of customer privacy and losses of customer data
G4-PR9	GRI 419	419-1	Non-compliance with laws and regulations in the social and economic area

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 4
Examining the Evolution of E-Government Development of Nations Through Machine Learning Techniques

Niguissie Mengesha
https://orcid.org/0000-0001-5673-5733
Brock University, Canada

Anteneh Ayanso
Brock University, Canada

ABSTRACT

Several initiatives have tried to measure the efforts nations have made towards developing e-government. The UN E-Government Development Index (EGDI) is the only global report that ranks and classifies the UN Member States into four categories based on a weighted average of normalized scores on online service, telecom infrastructure, and human capital. The authors argue that the EGDI fails in showing the efforts of nations over time and in informing nations and policymakers as to what and from whom to draw policy lessons. Using the UN EGDI data from 2008 to 2020, they profile the UN Member States and show the relevance of machine learning techniques in addressing these issues. They examine the resulting cluster profiles in terms of theoretical perspectives in the literature and derive policy insights from the different groupings of nations and their evolution over time. Finally, they discuss the policy implications of the proposed methodology and the insights obtained.

INTRODUCTION

The relevance of digitization and electronic service has become evident more than ever before because of the COVID-19 pandemic. The pandemic caused nations of the world to lockdown citizens, close economies, encourage and enforce social distancing measures, and discourage physical contact. To as-

DOI: 10.4018/978-1-7998-6985-6.ch004

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

sure existence and continuity of service, and fulfill the needs of customers, governments and enterprises aggressively worked on digitizing business processes and authority structures. Governments at all levels from city administration to regions through provinces to national governments sought to introduce information technology solutions and digitize processes, authority structures, and organizations so that employees could work from home/remote, and customers continue doing business from remote. Some mention COVID-19 as the single most powerful force that dictated governments and enterprises to digitize businesses and services and come up with digital strategies unlike any other force prior to that.

Electronic government (e-government for short) is the application of Information Technology (IT) to government services and procedures in order to deliver public services and information to citizens, businesses and governments electronically (Kim, Pan, & Pan, 2007). E-government redefines the interaction between administrations and customers (such as citizens, businesses and other administrations), and creates an electronic, minimal, more transparent, agile and accountable state (Ciborra, 2005). E-government has several potential benefits such as it increases accessibility of government services, allows greater public access to information, and makes governments more accountable to citizens (Hafkin, 2009; Verkijika & De Wet, 2018). It has also the potential to improve the quality and efficiency of processes, and reduce administrative burden on citizens, businesses and other administrations (Heeks, 2002).

Nations of the world have invested on IT infrastructure and human resource development, restructured and digitized work processes and structures, put in place necessary regulatory frameworks, and set priorities. Several studies have pointed out factors that determine the success and maturity of e-government (Das, Singh, & Joseph, 2017; Lee, Tan, & Trimi, 2005), while others tried to measure and benchmark e-government development efforts of nations using several benchmarking and ranking indices and models (Máchová & Lnenicka, 2015). Some of these models particularly focus on benchmarking online services or websites (Rorissa & Demissie, 2010; Verkijika & De Wet, 2018; West, 2004a), follow a more generalized approach with a focus on processes and services (Kunstelj & Vintar, 2004), and the United Nations e-government index, for example, uses a more broader model that includes tools to measure infrastructure, education and online services (UNDESA, 2020).

These evaluations and researches focus on the construction of e-government indices with a primary focus on ranking nations. While these evaluations and indices are vital to discover and communicate the current state and rank of each nation's e-government status (Kunstelj & Vintar, 2004), and serves as a good benchmark for nations to track their ranked positions from time to time, a mere increase or decrease of a nation's rank by few positions does not provide a bigger picture on their evolution towards making a major transformation over time (Das et al., 2017). A nation's attempt to bring major transformations need to be informed by its past and current profiles and the movement from one comparison set to another. In other words, if there is a major disparity in e-government status among groups of nations, it is important for a nation to set its goal based on this disparity and strive to advance to a high-achieving comparison set.

Secondly, each ranking model or approach has its own drawback. For example, studies show that the current UNDESA e-government ranking system may not clearly indicate the ranks of nations (Whitmore, 2012). The indices developed from the raw data seem to be artificial according to this study. The UNDESA index also places equal explanatory power to the three sub-indices such as IT infrastructure, human capital, and online service and takes the average of the three to derive e-government index for nations. Even if this assumption doesn't bring major change in the ranking, it has a negative impact when grouping countries together. UNDESA, for example, groups countries together as low, middle, high, and very high EGDI if the EGDI of a nation was 0 to 0.25, 0.25 to 0.5, 0.5 to 0.75 and from 0.75 to 1 respectively. We argue that this grouping may not bring countries with similar characteristics together,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

doesn't show the movement of nations over time, and does not help policy makers to learn from the experiences of other exemplary nations.

This chapter bridges these gaps by coming up with a model that classifies nations together according their e-government profiles (based on the UNDESA data) and identifying exemplary nations from each cluster. A nation makes a considerable progress only when it moves from one cluster to the other. As such, a nation can easily identify an exemplary nation to adapt exemplary theory and practice from, while formulating policies and plans. In this paper we employ cluster analysis to profile the e-government development status of the UN member states and their evolution over time from the 2008 to 2020 using the UNDESA e-government development index components. Based on the underlying structure of the data for each year, we attempt to capture the grouping of the nations as well as their movement from one group to another to identify meaningful profiles and extend the UNDESA rankings with additional insights and policy implications. By doing so, the research extends the methodology used in the UNDESA project and other scientific results presented in previous research to measure the e-government development status of nations.

Our focus in this paper is the UNDESA project and its subsequent indices as the approach and its data are more comprehensive and exhaustive than others, and it is easy to understand and easy to reproduce (Whitmore, 2012). The UNDESA project primarily assesses the extent to which governments or economies are equipped to deliver various governmental services online and exploit ICT for internal functioning of the government (Ahmed & Hussein, 2006; UNDESA, 2020). It has been ranking the UN Member States since 2001 by assigning an e-government index, which is calculated by averaging the normalized online service, human capital and telecommunication infrastructure indices (UNDESA, 2020). Using the key components of the UNDESA e-government index and unsupervised machine learning techniques, we analyze the e-government status profiles of the UN Member States to provide a rigorous quantitative assessment of their position and evolution over time during the period 2008-2020. We also explore any disparity in their e-government status profiles, typical characteristics, similarities and differences, and their evolution over time.

The remainder of this paper is organized as follows: we begin by providing a review of the relevant studies that have contributed to this body of research. This will be followed by an overview of the UNDESA survey data which is the basis of our data analysis and the methodology used in this research. The subsequent section presents the analyses and the results obtained. Finally, we discuss the policy implications of our findings and the contributions to the literature.

E-GOVERNMENT DEVELOPMENT EVALUATIONS: WHAT AND HOW TO MEASURE IT

The purpose of e-government is to deliver government information and services to citizens, government and non-government organizations, and businesses using the Internet and World Wide Web technologies (Das et al., 2017; UNDESA, 2001). E-government redefines the interaction between administration and citizens and organizations, and creates an electronic, minimal, more transparent, agile, and accountable state (Ciborra, 2005). It has the potential to improve the quality and efficiency of processes and reduce administrative burden on citizens by making their interactions with public administrations faster and efficient, more convenient and transparent, and less costly (Badri & Alshare, 2008; OECD, 2005; Prahono & Elidjen, 2015). To achieve these goals, all levels of governments are required to restructure, redesign,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

and digitize processes, authority structures, data, and rules and regulations, and deploy the required IT infrastructure. The second major component entails for connecting this redesigned and restructured system with citizens and organizations that do businesses with the government to deliver information and services. The system put in place should allow users to access, view, or perform operational functions; available whenever required; available to operate on different devices or software platforms; function correctly and provides accurate information; and it has to be easy to learn and efficient and satisfying to use (Das et al., 2017). Once such systems are put in place, with proper training and awareness interested and able citizens and organizations would use the system.

Studies have identified several factors that determine the diffusion and success of e-government. For example, a country's social, political, and economic compositions as key factors determine a country's e-government development program (UNDESA, 2001). The political environment and national policies, more importantly, whether a government is open to political reform or not, are also among the factors that determine the success of e-government (Lee et al., 2005). A state's telecommunications infrastructure, the strength of its human capital, innovative capacity, political will and commitment of the national leadership, wealth, rule of law, and policy and administrative priorities are also among the factors that play important roles in the development and implementation of e-government programs (Ifinedo, 2012; UNDESA, 2001). Our focus in this chapter is not about identifying factors that influence e-government development but on the measurement tools and models suggested and applied to assess, measure, and benchmark e-governance capabilities of nations. We also identify the attributes of e-government that these models measure.

There has been a large pool of initiatives focusing on measurements and benchmarks for e-government readiness (Ayanso, Chatterjee, & Cho, 2011). These include public-private initiatives, typically commissioned by governmental or inter-governmental bodies and executed by private consultants (e.g., (Capgemini, 2005, 2006)); private initiatives that were developed to serve policy makers as well as private investors interested in e-governance (Accenture, 2006); academic efforts carried out by the Center for Public Policy at Brown University (West, 2003, 2004a, 2004b, 2005, 2006, 2007); and finally, the United Nations e-government surveys, which are conducted annually since 2001, to assess e-government readiness of nations (UNDESA, 2001, 2003, 2004, 2005, 2008, 2010, 2018, 2020). Although comprehensive reviews of these benchmarking efforts are provided by other researchers (Bannister, 2007; Jansen, 2006; Kunstelj & Vintar, 2004), we point out the main themes of these models here.

The above benchmarking and measurement models and tools can be classified into two categories: those attempts that study government websites and those that analyze a country's capability by surveying government websites and considering additional indicators such as human capital and IT infrastructure. For example, West, in the annual reports issued from 2002 to 2007, analyzed national government websites belonging to executive offices (such as a president, prime minister, ruler, party leader, or royalty), legislative offices (such as Congress, Parliament, or People's Assemblies), judicial offices (such as major national courts), Cabinet offices, and major agencies serving crucial functions of government, such as health, human services, taxation, education, interior, economic development, administration, natural resources, foreign affairs, foreign investment, transportation, military, tourism, and business regulation (West, 2003, 2004b, 2005, 2006, 2007). These reports evaluated websites for the presence or absence of features and functions dealing with information availability, accessibility, security features, service delivery, digital signature, and personalization. The result is aggregated to show the maturity of websites and, hence, rank nations. Similarly, other studies and frameworks (Affisco John & Soliman Khalid, 2006; Al-adawi, Cardiff, & Pallister, 2005; Rorissa & Demissie, 2010) attempted to measure and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

rank the maturity of electronic delivery of government services by analyzing the features and functions of government websites and classifying them into four stages/levels of development of e-Government services such as publishing (web presence), interacting, transacting, and transforming (integration). Studies analyzed the maturity of government websites at a given point in time, like the ones mentioned above, or over a period of time (Das et al., 2017; Ifinedo, 2012).

The focus on website disregards other important factors that may impact the use and success of e-government initiates. The existence of a matured government website is not a guarantee to its utilization. The use and consequence of a technology emerge unpredictably from the complex interactions between social and technical factors. Therefore, considering the social component and the infrastructural component gives a better picture to understand e-government initiatives. To this end, the second category like the United Nations E-Government Survey, measure performance of countries based on the adequacy of telecommunication infrastructure, the ability of human resources to promote and use ICTs, and the availability of online services and content (UNDESA, 2020). The United Nations Department of Economic and Social Affairs (UN DESA) has been publishing the Survey since 2001 every other year; it has published 10 reports so far. The Survey tracks progress of e-government development via the United Nations E-Government Development Index (EGDI). The EGDI assesses the e-government development of a nation and it is a composite index based on the weighted average of three normalized indices: Telecommunications Infrastructure Index (TII), Human Capital Index (HCI), and Online Service Index (OSI) (UNDESA, 2020). The Survey is the only global report that assesses the e-government development status of all United Nations Member States. Because of availability, recency, coverage, and relevance, and following prior studies (Ayanso et al., 2011; Mengesha, Ayanso, & Demissie, 2020), we used the UNDESA's EGDI data to profile and examine the e-government development status of UN Member States from 2008 to 2020. The next section describes the data and methods in detail.

DATA AND METHODOLOGY

Data

In this study, we focus on the underlying components of the UN EGDI data collected by the UN DESA. Unlike other indices, the UN EGDI index and its components data has a broader coverage and most current reporting of e-government assessment, and the data is publicly available for free. This makes replication, critique, and extension of our results possible and the broad coverage also assures that our findings are truly generalizable and free from selection-related biases.

The EGDI is a composite index based on the weighted average of the normalized Telecommunications Infrastructure Index (TII), the Human Capital Index (HCI), and Online Service Index (OSI). The TII represents the physical information technology infrastructure based on data provided by the International Telecommunications Union (ITU). The HCI measures the general literacy and educational levels of the citizens based on data mainly provided by the United Nations Educational, Scientific and Cultural Organization (UNESCO). The OSI measures the presence or absence of comparable online functionalities grouped hierarchically into *emerging, enhanced, interactive, transactional*, and *connected* stages of web-based service delivery based on UNDESA's assessment of national online presence of UN Member States. Despite some changes in the compositions of some of its components, the UN EGDI assessment has a more coherent reporting every other year, particularly since 2008. For example, in the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

2005 E-government Readiness Report (UNDESA, 2005), the telecommunication infrastructure index had six components (i.e., PCs, Internet users, Telephone lines, Online population, Mobile subscribers, and TVs). In the 2008 UN E-government Survey (UNDESA, 2008), the TII had only five components by replacing online population and TV with Broadband. The current TII component has four components such as Internet users, active mobile-broadband subscriptions, mobile subscribers, and fixed broadband subscriptions (UNDESA, 2020). The components of HCI remained stable over the years except the inclusion of expected years of schooling and mean years of schooling in 2014. Prior to 2014, the HCI was computed by using adult literacy rate and the combined primary, secondary and tertiary gross enrolment ratio (UNDESA, 2020).

In our study, we used the data from 2008 survey onwards. Since then, the index and its key components have remained relatively consistent and coherent. Using the UN EGDI data set, we assembled data for 175 countries out of the 193 UN Member States for which complete data set for the components was available for the period from 2008 to 2020. Our goal in this paper is not to propose a new way of constructing the e-government index but to assess the profiles and evolution of UN Member States based on the key components of the UN EGDI. Table 1 presents summary statistics for the three components for the period 2008-2020.

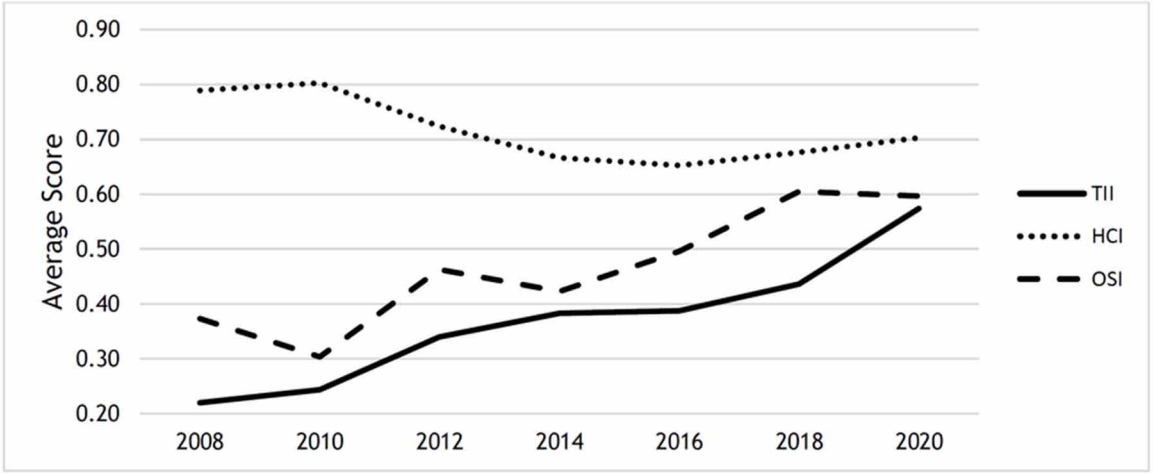
Table 1. Summary Statistics of E-government Index Components

Year	Telecommunications Infrastructure Index				Human Capital Index				Online Service Index			
	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
2008	0.22	0.15	0.00	0.81	0.79	0.86	0.25	0.99	0.37	0.35	0.01	1.00
2010	0.24	0.18	0.01	0.77	0.80	0.88	0.28	1.00	0.30	0.28	0.00	1.00
2012	0.34	0.28	0.01	1.00	0.72	0.79	0.11	1.00	0.46	0.44	0.08	1.00
2014	0.38	0.35	0.02	0.94	0.67	0.71	0.12	1.00	0.42	0.39	0.01	1.00
2016	0.39	0.37	0.03	0.85	0.65	0.70	0.05	1.00	0.50	0.50	0.02	1.00
2018	0.44	0.43	0.05	0.86	0.68	0.72	0.09	1.00	0.60	0.65	0.05	1.00
2020	0.57	0.59	0.07	1.00	0.70	0.74	0.13	1.00	0.60	0.62	0.03	1.00

From the summary statistics and as shown in Figure 1, we can observe that the telecommunication infrastructure component has shown incremental progress in each reporting year since 2008. On the other hand, the human capital index component has shown a declining pattern since 2010 on average. The drop in HCI in 2014 can be attributed to the inclusion of expected years of schooling and mean years of schooling to measure human capital besides adult literacy and gross enrollment ratio. There has been also a sign of improvement on the online service delivery component since 2014. Overall, these summary statistics show a tendency to focus on the infrastructure domain of e-government readiness. In the next section, we examine these profiles and the changes over time more rigorously based on segmentation or cluster analysis on the three components of the e-government index.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Average E-government Index Components (2008 to 2020)



Analysis Method

We find the method used in EGDI is intuitive and easy to follow. However, we believe constructive refinements can be applied to improve upon several aspects of the methodology. First, the simplistic grouping of the countries can be improved with a more rigorous technique that could help nations not only monitor ranking but also draw lessons from peer nations. We believe the use of clustering and the depiction of a nation's membership profile and evolution over time allows policy makers and scholars assess e-government efforts and progress in a more constructive context. In other words, the cluster-based methodology allows nations monitor their progress through the lenses of the peers they belong with and the peers they have surpassed them over time. We achieve this by capturing the natural grouping of the countries, identify meaningful profiles in terms of the different dimensions of the data, as well as the movement of the countries from one group to another based on the underlying structure of the data. Closing the e-Government gap by a wider margin and leapfrogging from one state of membership profile to a more advanced state of membership profile is not only aspirational for nations but also instructional from a policy standpoint. When nations leapfrog to an advanced club of other nations, the path they follow can be easily portrayed through the cluster profile, along with the policy implications of the specific improvements on the key dimensions of e-Government.

Although the use of machine learning techniques is becoming very common in many research domains, their applications in the e-Government domain has been very limited. Our use of unsupervised machine learning technique or clustering will provide a unique perspective and contribution to the current e-Government literature. Cluster analysis is a commonly used data mining technique to organize observed data into meaningful groups (or clusters). More specifically, clustering as a data analytics technique is a process of identifying a set of individuals, nations, or objects in such a way that those in the same group are more similar to each other than to those in other groups. The method has been applied to a wide variety of problem contexts in private and public sectors. Examples include identifying market segments (Chaturvedi, Carroll, Green, & Rotondo, 1997; Dolnicar, 2003), or profiling mobile Internet adopters (Okazaki, 2006). More importantly, past studies that are directly relevant to this research include profiling the level of access and use of ICT infrastructures such as PCs, Internet, telephone lines, and mobile

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

phones (Ayanso, Cho, & Lertwachara, 2010). Ayanso et al. (2014) also extended this research using the more comprehensive ICT Development Index data set, which is constructed based on the three key components of the information society model (i.e., Access, Use, and Skills). Their extended analyses incorporated not only the access to ICT infrastructures, but also ICT usage and skills to examine the progress in ICT adoption made by developing countries in each global region over a five-year span (2002-2007). Another study that utilized cluster analysis examined the levels of the ICT infrastructure development among the European Union countries (Vicente Cuervo & López Menéndez, 2006). More recently, Mengesha, Ayanso, and Demissie (2020) analyzed and profiled the e-government development status of African nations.

The specific cluster analysis technique we use in this paper is the Expectation-Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977). The EM algorithm is a distribution-based clustering algorithm which relies on a maximum likelihood computation. The method is based on the assumption that the indicators variables that are used to generate clusters follow one of several statistical distributions. Thus, the main goal in the algorithm's routine is to identify the parameters of each distribution as well as their values (Jain, Murty, & Flynn, 1999). For this study and analysis, we choose the EM algorithm for several important reasons. First, unlike other clustering methods, the EM method is particularly suggested to deal with situations where there are data quality issues such as missing values in one or more variables. The second advantage of the EM algorithm is that it provides flexibility for the user in experimenting with an initial number of desired clusters or to let the algorithm suggest the appropriate number of clusters based on the fit to the underlying data and a statistical criterion. In the auto-clustering option, the EM algorithm selects the number of clusters by maximizing the log likelihood of future data through cross-validation. The availability of these features has allowed us to explore the research data and determine an appropriate number of segments in the underlying data. Finally, given that the EM algorithm is readily available in many machine learning as well as statistical software packages, it makes it easier for the replication as well as extension of our analysis in future research. While the EM algorithm can handle mixed variables (i.e., numeric and nominal variables), our current research includes only numerical variables. However, this is another advantage of the EM algorithm for future research that may include other variables and different forms of measurement scales.

In this paper, we use the EM algorithm implementation in one of the popular data mining tools, namely, Weka (Hall et al., 2009; Witten & Frank, 2002). According to the EM algorithm in Weka, a probability distribution is assigned to determine the probability of a given data point belonging to a specific cluster. As mentioned before, the user may specify *apriori* how many clusters to generate or the algorithm can auto-determine the number of clusters through a cross validation procedure. In the first step, all data points are assigned into one cluster and the training data is randomly split into n folds. For each fold, the procedure involves computing a log likelihood value, followed by an average of the different folds. If the computed average has increased from the last iteration, then the number of clusters is increased by one and this process is repeated (Hall et al., 2009; Witten & Frank, 2002).

E-GOVERNMENT PROFILE AND EVOLUTION OF NATIONS

The e-government data used in this research includes the years 2008 through 2020. Since the UN EGDI component values are already given in a normalized scale, standardization of the data points was not necessary. Using the EM algorithm in Weka, we input the UN EGDI components – online service, human

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

capital, and telecommunications infrastructure indices – of the 175 UN Member States for unsupervised clustering. For the sake of comparison and to facilitate further analysis of the result in relation with the economic status of nations, we set the resulting number of clusters to 4, i.e., we set Weka to cluster the data into 4 clusters. We run the clustering algorithm seven times, one each for each year, and identified the clusters according to each cluster's centroids. Following the presentation format of prior related studies in the ICT domain (Ayanso, Cho, & Lertwachara, 2014; Ayanso & Lertwachara, 2015), we labeled the four clusters of e-government profiles of the UN Member States as "Beginner", "Emergent", "Progressor", and "Veteran". The "Beginners" are novice and learners in e-government development; the "Emergent" are up-and-coming and growing; the "Progressors" are those that are progressing and expanding; and the "Veterans" are proficient and connoisseurs in their e-government development. The following table, Table 2, lists the centroids of each cluster for each EGDI component and the number of nations in each cluster for each year from 2008 to 2020.

Table 2. E-government Cluster Centroids and Number of Nations (#Nat) in Each Cluster – Beginner, Emergent, Progressor, and Veteran (2008-2020)

Year	Beginner				Emergent				Progressor				Veteran			
	OSI	HCI	TII	#Nat	OSI	HCI	TII	#Nat	OSI	HCI	TII	#Nat	OSI	HCI	TII	#Nat
2008	0.17	0.57	0.03	54	0.32	0.81	0.11	33	0.42	0.88	0.25	57	0.67	0.97	0.61	31
2010	0.13	0.58	0.04	51	0.25	0.83	0.13	36	0.33	0.89	0.29	57	0.60	0.97	0.61	31
2012	0.25	0.48	0.08	51	0.41	0.73	0.21	33	0.51	0.82	0.41	56	0.75	0.92	0.73	35
2014	0.19	0.44	0.12	56	0.41	0.71	0.34	64	0.56	0.81	0.60	31	0.81	0.90	0.80	24
2016	0.25	0.44	0.15	62	0.50	0.70	0.39	61	0.68	0.83	0.58	23	0.84	0.86	0.75	29
2018	0.37	0.47	0.18	61	0.60	0.71	0.42	46	0.74	0.80	0.60	34	0.90	0.87	0.76	34
2020	0.34	0.45	0.23	40	0.45	0.64	0.48	38	0.70	0.77	0.66	52	0.84	0.90	0.86	45

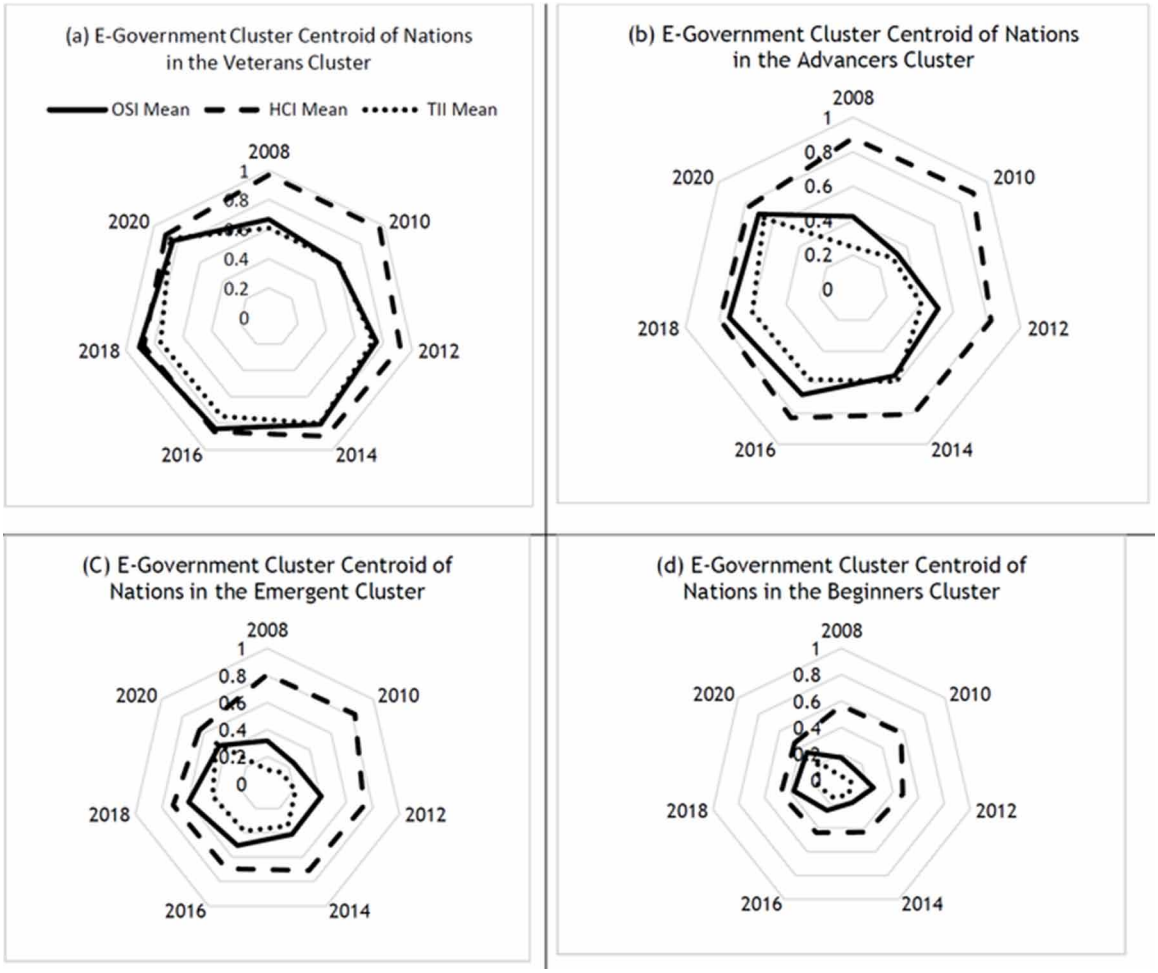
The following chart, Figure 2, illustrates the centroids of each cluster for each EGDI component for each year from 2008 to 2020. As we can see from the figure, there is a pronounced difference among the clusters and between Veterans and Beginners in each EGDI component. There is a marked increase in each component over the years for each cluster, and the gap between the three components has been narrowing down over time. The human capital component is the highest for each cluster followed by online service and telecom infrastructure. The result also shows the need for each nation in each cluster to put more effort and resource on telecom infrastructure development followed by online services expansion.

As shown in Figure 3, only a handful of countries are in the top rank, 19% (33 out of 175) on average and the size remained the same over the years (18% (31 out of 175) in 2008 and 17% (29 out 175) in 2016), and (26% (46 out of 175) in 2020. Slight shifts happed in 2012, 2014, and 2020. The number of countries in the Veteran cluster rose to 20% (35 out of 175) in 2012, fell by 6% from the previous year to 14% (24 out of 175) in 2014, and jumped from 19% to 26% in 2020.

Over the years (from 2008 to 2020), several nations fell to the Emergent cluster from Progressor. In 2008, 33% of the nations (57 out of 175) were in Progressor cluster and that number fell to 13% (23 out of 175) in 2016. Major such change happened in 2014. The number of countries in the Progressor cluster was 32% (56 out of 175) in 2012 and that number dropped to 18% (31 out of 175) in 2014. That

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. E-government Cluster Centroids of Nations (2008-2020) – (a) Veteran, (b) Progressor, (c), Emergent, and (d) Beginner clusters

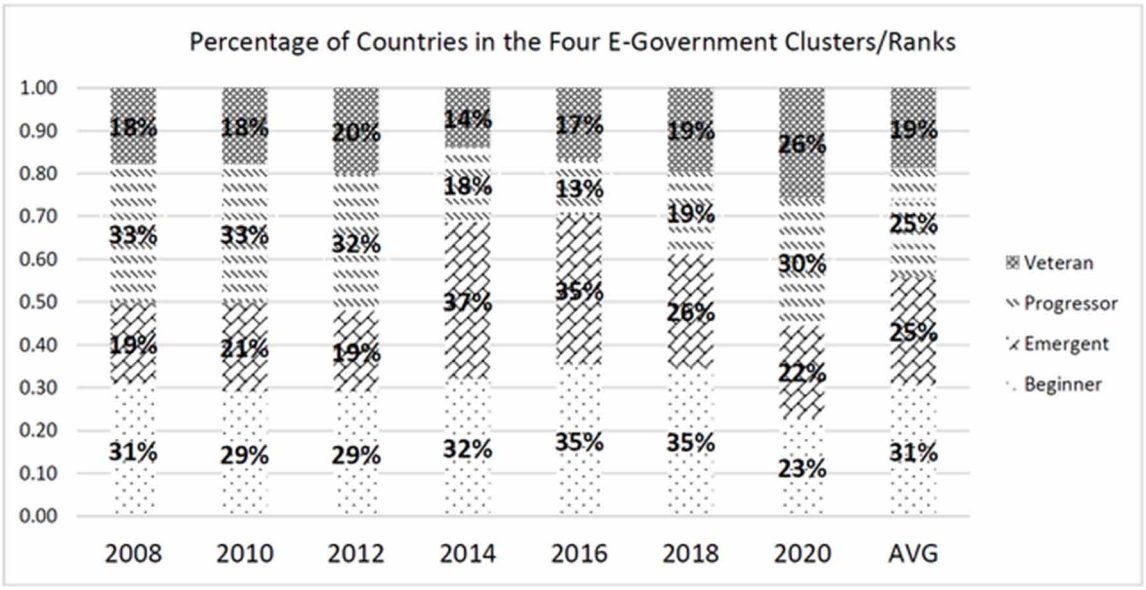


accounts for a 16% change. The trend changes in 2018 and 2020. In 2018, many countries (9%) joined either the Progressor or the Veteran cluster from the Emergent cluster. Several countries moved to the next upper level in 2020. The number of countries in the Beginner cluster fell by 8% from 2018 to 2020; the number of countries in the Veteran and Progressor clusters rose by 18% from 38% in 2018 to 56% in 2020; and those in the Beginner and Emergent clusters dropped by 16% from 61% in 2018 to 61% in 2020. Although that is the case, relatively the number of countries in the Beginner cluster remained the same. The number of countries in the Beginner and Emergent categories were very high, 69% and 70% in 2014 and 2016 respectively. The number drops to 43% in 2020.

Studies have shown that wealth or economic status of a nation is an important factor that determines the maturity of e-government (Das et al., 2017; Ifinedo, 2012). We have done a country level analysis to learn about this relationship, profile each nation, and understand their evolution over time. The complete result of the analysis is presented in Table 5 as an appendix. Our country level analysis tells two different stories: some countries performed exceptionally well even if they are not wealthy while wealthy

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Percentage of Countries in the Four E-Government Clusters from 2009-2020



nations did not. Several countries, about 18 (approximately 32% of the nation) from the High-Income Group according to the World Bank classification of countries for the 2021 fiscal year (World Bank, 2020), did not perform well in e-government development despite their high economic status. These countries are supposed to be in the Veteran cluster, but they are either in Progressor or Emergent clusters together with Upper and Lower Middle-Income countries. Except Antigua and Barbuda, Barbados, Croatia, Hungary, Kuwait, and United Arab Emirates, which were in the Veteran cluster one year or for a maximum of three years, none of these countries were able to enjoy e-government services like their economic counterparts. The following table, Table 3, shows the list of countries that did not perform well in e-government development in accordance with their economic status. As shown in the table, several countries (22 about 43%) in the Upper Middle-Income class are in the Emergent or Beginner clusters while they are supposed to be in the Progressor cluster. The same is true for Lower Middle-Income countries; 16 countries (about 36%) from the Low Middle-Income group did not perform well as expected as shown in the table.

The following table, Table 4, lists exemplary countries in e-government development from the Upper Middle-Income, Middle-Lower Income, and Low Income countries. These nations were performing very well in e-government development despite their economic status. Although not year after year, Argentina, Belarus, Bulgaria, Kazakhstan, and Russian Federation, which accounts for 10% of the nations from the Upper Middle-Income countries joined the Veteran cluster together with High Income countries. Bolivia, Cape Verde, Kyrgyzstan, Mongolia, Philippines, Republic of Moldova, Sri Lanka, Tunisia, Ukraine, Uzbekistan, and Vietnam, which accounts for a quarter of the nations in the Lower Middle-Income group, are also among the nations that are performing well in e-government development. Syrian Arab Republic and Tajikistan, accounting for 9% of the countries, have been doing well from the Low-Income group.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Countries that did not Perform Well in E-Government Development

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
1	Andorra	HIGH INCOME	Progressor	Progressor	Progressor	Progressor	Emergent	Progressor	Progressor
2	Antigua and Barbuda		Progressor	Progressor	Veteran	Progressor	Emergent	Emergent	Progressor
5	Bahamas		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
7	Barbados		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Progressor
9	Brunei Darussalam		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
12	Croatia		Progressor	Progressor	Veteran	Progressor	Progressor	Progressor	Progressor
21	Hungary		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Progressor
27	Kuwait		Progressor	Progressor	Progressor	Progressor	Veteran	Progressor	Progressor
33	Mauritius		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
37	Oman		Emergent	Emergent	Progressor	Emergent	Emergent	Progressor	Progressor
38	Panama		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
41	Qatar		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor
43	Romania		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
44	Saint Kitts and Nevis		Progressor	Progressor	Progressor	Progressor	Emergent	Progressor	Progressor
46	Seychelles		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
48	Slovakia		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor
53	Trinidad and Tobago		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
54	United Arab Emirates		Progressor	Progressor	Progressor	Progressor	Veteran	Veteran	Progressor
63	Belize	UPPER MIDDLE INCOME	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
65	Botswana		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
71	Cuba		Beginner	Emergent	Beginner	Beginner	Beginner	Beginner	Beginner
72	Dominica		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
75	Equatorial Guinea		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
77	Gabon		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
79	Grenada		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
80	Guatemala		Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
81	Guyana		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Emergent
84	Iraq		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
85	Jamaica		Progressor	Progressor	Emergent	Emergent	Emergent	Emergent	Emergent
86	Jordan		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
88	Lebanon		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
90	Maldives		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
93	Namibia		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Emergent
98	Saint Lucia		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
99	Saint Vincent and the Grenadines		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
100	Samoa		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner
103	Suriname		Emergent	Emergent	Progressor	Emergent	Emergent	Emergent	Emergent
105	Tonga		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
107	Turkmenistan		Beginner	Beginner	Beginner	Emergent	Beginner	Beginner	Emergent
108	Venezuela		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
110	Angola	LOWER MIDDLE INCOME	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
112	Benin		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
116	Cameroon		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
118	Comoros		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
119	Congo		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
121	Djibouti		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
125	Honduras		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner
129	Lao Peoples Democratic Republic		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
131	Mauritania		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
136	Nigeria		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
137	Pakistan		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
138	Papua New Guinea		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
141	Sao Tome and Principe		Emergent	Emergent	Beginner	Beginner	Beginner	Beginner	Beginner
142	Senegal		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
143	Solomon Islands		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
148	United Republic of Tanzania		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Countries that are Performing Well in E-Government Development

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
59	Argentina	UPPER MIDDLE INCOME	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
62	Belarus		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
67	Bulgaria		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
87	Kazakhstan		Emergent	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
97	Russian Federation		Progressor	Progressor	Veteran	Progressor	Progressor	Veteran	Veteran
114	Bolivia		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
117	Cape Verde	LOWER MIDDLE INCOME	Emergent	Emergent	Emergent	Emergent	Emergent	Veteran	Veteran
128	Kyrgyzstan		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
132	Mongolia		Emergent	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
139	Philippines		Progressor	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
140	Republic of Moldova		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
144	Sri Lanka		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
146	Tunisia		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
147	Ukraine		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
149	Uzbekistan		Beginner	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
151	Vietnam		Emergent	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
171	Syrian Arab Republic	LOW INCOME	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
172	Tajikistan		Beginner	Beginner	Emergent	Emergent	Beginner	Beginner	Emergent

DISCUSSION

The analysis of the EGDI data of the UN Member States using unsupervised clustering technique adds one important dimension to understanding the current state and the progress nations have made in realizing the full potential of e-government over the years. Our analysis resulted in four groups – Beginner, Emergent, Progressor, and Veteran. The number of nations in each group changed over the years as policy directions and resources that nations put in online service, human capital, and telecommunications development changed. The movement of nations from one group to the other indicates performance relative to other nations. Those nations that join the upper group from the lower indicates positive performance, while dropping down to the lower group indicates otherwise. Unlike the simple and structured UNDESA grouping of nations, the clustering approach used in this chapter considers the likelihood of nations to fall in a group based on their EGDI components.

Apart from learning where nations stand in relation to others based on a complex grouping algorithm, our analysis shows the trajectory of nations in e-government development over the years. Kazakhstan from upper middle-income economic group, for example, was in the Emergent group in 2008, joined the next upper group, the Progressor, in 2010 and remained in the same group until 2020. In 2020, Kazakhstan joined the Veteran group showing a very good progress. Vietnam from lower middle-income group was in the Emergent group in 2008, joined the Progressor group in 2010 and remained in the same group

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

until 2014. In 2014, Vietnam dropped to the Emergent group and remained in the same group until 2018. Vietnam joined the next upper group, the Progressor, in 2020.

This type of longitudinal study is sought after but lacking (Das et al., 2017) in the literature with a few exceptions (Das et al., 2017; Ifinedo, 2012). Most empirical researches on e-government have been dominated by cross-sectional analysis (Das et al., 2017). The historical trajectory, apart from showing the movement of nations, paves the way for policy and decision makers to explore and identify policies and related issues that contributed to the upward or downward movement. For example, a nation like Vietnam, can ask why and how they moved from the Emergent to the Progressor group in 2010 and back to the Emergent group in 2014. This type of analysis helps a nation to see inward and rectify or strengthen policy direction or resource allocation accordingly. It also permits nations to see the trajectories of others and learn from their successes or failures.

Wealth or economic status is among the dominant factors that determine the maturity of a nation's e-government development (Das et al., 2017; Ifinedo, 2012). To this end, our analysis uncovered two important findings: (1) some nations performed exceptionally well in e-government development despite their economic grouping; and (2) some other nations performed below the expectation. Although not consistently, Argentina, Belarus, Bulgaria, Kazakhstan, and Russian Federation from the Upper Middle-Income countries joined the Veteran group together with High Income countries. Bolivia, Cape Verde, Kyrgyzstan, Mongolia, Philippines, Republic of Moldova, Sri Lanka, Tunisia, Ukraine, Uzbekistan, and Vietnam from the Lower Middle-Income group joined the Progressor group together with Upper Middle-Income countries. The movements of these nations are remarkable and offers an opportunity for other nations to learn from these successful nations. Studying the policies and actions of these successful nations regarding e-government development offers an opportunity for other similar nations to learn from the successes of these nations.

Some countries did not perform well in e-government development despite their high economic status. As indicated in Table 3, several nations from the High-Income group were either in the Progressor or Emergent group together with Upper and Lower Middle-Income nations. The same applies to some nations in the Upper and Lower Middle-Income group. These nations could learn by looking inwards or outward by the studying the policies and actions of successful nations.

CONCLUSION

In this chapter, we analyzed and profiled the UN Member Nations e-government development status using the UN EGDI online service, human capital, and telecommunications indices and the Expectation Maximization unsupervised machine learning algorithm in Weka. The algorithm clustered the 175 Member States into four clusters, and we named the clusters as Beginner, Emergent, Progressor, and Veteran in an increasing order of maturity in e-government development. Our longitudinal study reveals the evolution and current state of e-government development status of the UN Member Nations relative to others. Some nations performed exceptionally well despite their economic status while some wealthy nations did not perform as expected. These groupings and the movement from one group to the other provide opportunities for Nations to learn from exemplary Nations in their effort to fine tune policies and actions in e-government development. Our approach, analysis, and the identification of exemplary nations from each group and over the years, helps policy and decision makers where to look for policy and guidance. The chapter also demonstrates the relevance of machine learning approach, particularly

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the unsupervised clustering technique, in classifying nations and thereby refining the current UNDESA approach.

REFERENCES

Accenture. (2006). *E-government leadership: Building the trust*. Accenture.

Affisco John, F., & Soliman Khalid, S. (2006). E-government: A strategic operations management framework for service delivery. *Business Process Management Journal*, *12*(1), 13–21. doi:10.1108/14637150610643724

Ahmed, A.-O., & Hussein, A.-O. (2006). E-government readiness assessment model. *Journal of Computational Science*, *2*(11), 841–845. Advance online publication. doi:10.3844/jcssp.2006.841.845

Al-adawi, Z., Cardiff, Y., & Pallister, J. (2005). *Conceptual Model of Citizen Adoption of E-Government.* Paper presented at the Second International Conference on Innovations in Information Technology (IIT'05). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.2165&rep=rep1&type=pdf

Ayanso, A., Chatterjee, D., & Cho, D. I. (2011). E-government readiness index: A methodology and analysis. *Government Information Quarterly*, *28*(4), 522–532. doi:10.1016/j.giq.2011.02.004

Ayanso, A., Cho, D. I., & Lertwachara, K. (2010). The digital divide: Global and regional ict leaders and followers. *Information Technology for Development*, *16*(4), 304–319. doi:10.1080/02681102.2010.504698

Ayanso, A., Cho, D. I., & Lertwachara, K. (2014). Information and communications technology development and the digital divide: A global and regional assessment. *Information Technology for Development*, *20*(1), 60–77. doi:10.1080/02681102.2013.797378

Ayanso, A., & Lertwachara, K. (2015). An analytics approach to exploring the link between ict development and affordability. *Government Information Quarterly*, *32*(4), 389–398. doi:10.1016/j.giq.2015.09.009

Badri, M. A., & Alshare, K. (2008). A path analytic model and measurement of the business value of e-government: An international perspective. *International Journal of Information Management*, *28*(6), 524–535. doi:10.1016/j.ijinfomgt.2006.10.004

Bannister, F. (2007). The curse of the benchmark: An assessment of the validity and value of e-government comparisons. *International Review of Administrative Sciences*, *73*(2), 171–188. doi:10.1177/0020852307077959

Capgemini. (2005). *Online availability of public services: How is europe progressing?* Retrieved from http://www.umic.pt/images/stories/publicacoes/2004[1].01%20-%20Web-based%20Survey%20on%20Electronic%20Public%20Services%20-%20Fourth%20measurement.pdf

Capgemini. (2006). *Online availability of public services: How is europe progressing?* Retrieved from https://www.ut.is/media/Skyrslur/CapGemini_2006.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chaturvedi, A., Carroll, J. D., Green, P. E., & Rotondo, J. A. (1997). A feature-based approach to market segmentation via overlapping k-centroids clustering. *JMR, Journal of Marketing Research*, *34*(3), 370–377. doi:10.1177/002224379703400306

Ciborra, C. (2005). Interpreting e-government and development: Efficiency, transparency or governance at a distance? *Information Technology & People*, *18*(3), 260–279. doi:10.1108/09593840510615879

Das, A., Singh, H., & Joseph, D. (2017). A longitudinal study of e-government maturity. *Information & Management*, *54*(4), 415–426. doi:10.1016/j.im.2016.09.006

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B. Methodological*, *39*(1), 1–22. doi:10.1111/j.2517-6161.1977.tb01600.x

Dolnicar, S. (2003). *Using cluster analysis for market segmentation-typical misconceptions, established methodological weaknesses and some recommendations for improvement*. Academic Press.

Hafkin, N. J. (2009). *E-government in africa: An overview of progress made and challenges ahead*. Paper presented at the UNDESA/UNPAN workshop on electronic/mobile government in Africa: Building Capacity in Knowledge Management through Partnership, The United Nations Economic Commission for Africa.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explorations*, *11*(1), 10–18. doi:10.1145/1656274.1656278

Heeks, R. (2002). *Egovernment in Africa: Promise and practice*. Academic Press.

Ifinedo, P. (2012). Factors influencing e-government maturity in transition economies and developing countries: A longitudinal perspective. *SIGMIS Database*, *42*(4), 98–116. doi:10.1145/2096140.2096147

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323. doi:10.1145/331499.331504

Jansen, A. (2006). *Assessing e-government progress– Why and what.* https://www.uio.no/studier/emner/jus/afin/FINF4001/h05/undervisningsmateriale/AJJ-nokobit2005.pdf

Kim, H. J., Pan, G., & Pan, S. L. (2007). Managing it-enabled transformation in the public sector: A case study on e-government in south korea. *Government Information Quarterly*, *24*(2), 338–352. doi:10.1016/j.giq.2006.09.007

Kunstelj, M., & Vintar, M. (2004). Evaluating the progress of e-government development: A critical analysis. *Info. Pol., 9*(3,4), 131-148.

Lee, S. M., Tan, X., & Trimi, S. (2005). Current practices of leading e-government countries. *Communications of the ACM*, *48*(10), 99–104. doi:10.1145/1089107.1089112

Máchová, R., & Lnenicka, M. (2015). Reframing e-government development indices with respect to new trends in ict. *Review of Economic Perspectives*, *15*(4), 383–412. doi:10.1515/revecp-2015-0027

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mengesha, N., Ayanso, A., & Demissie, D. (2020). Profiles and Evolution of E-Government Readiness in Africa: A Segmentation Analysis. *International Journal of Information Systems and Social Change*, *11*(1), 43–65. doi:10.4018/IJISSC.2020010104

OECD. (2005). *E-government for better government*. OECD.

Okazaki, S. (2006). What do we know about mobile internet adopters? A cluster analysis. *Information & Management*, *43*(2), 127–141. doi:10.1016/j.im.2005.05.001

Prahono, A., & Elidjen. (2015). Evaluating the role e-government on public administration reform: Case of official city government websites in indonesia. *Procedia Computer Science*, *59*, 27–33. doi:10.1016/j.procs.2015.07.334

Rorissa, A., & Demissie, D. (2010). An analysis of african e-government service websites. *Government Information Quarterly*, *27*(2), 161–169. doi:10.1016/j.giq.2009.12.003

UNDESA. (2001). *Benchmarking e-government: A global perspective - assessing the progress of the un member states*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/English.pdf

UNDESA. (2003). *Un global e-government survey 2003*. Retrieved from https://publicadministration.un.org/egovkb/portals/egovkb/Documents/un/2003-Survey/unpan016066.pdf

UNDESA. (2004). *Un global e-government readiness report 2004: Towards access for opportunity*. Retrieved from https://publicadministration.un.org/egovkb/portals/egovkb/Documents/un/2004-Survey/Complete-Survey.pdf

UNDESA. (2005). *Un global e-government readiness report 2005: From e-government to e-inclusion* Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2005-Survey/Complete-survey.pdf

UNDESA. (2008). *United nations e-government survey 2008: From e-government to connected governance* Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2008-Survey/Complete-survey.pdf

UNDESA. (2010). *United nations e-government survey 2010: Leveraging e-government at a time of financial and economic crisis*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2010-Survey/Complete-survey.pdf

UNDESA. (2018). *United nations e-government survey 2018: Gearing e-government to support transformation towards sustainable and resilient societies*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2018-Survey/E-Government%20Survey%202018_FINAL%20for%20web.pdf

UNDESA. (2020). *E-government survey 2020: Digital government in the decade of action for sustainable development*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2020-Survey/2020%20UN%20E-Government%20Survey%20(Full%20Report).pdf

Verkijika, S. F., & De Wet, L. (2018). A usability assessment of e-government websites in sub-saharan africa. *International Journal of Information Management*, *39*, 20–29. doi:10.1016/j.ijinfomgt.2017.11.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vicente Cuervo, M. R., & López Menéndez, A. J. (2006). A multivariate framework for the analysis of the digital divide: Evidence for the european union-15. *Information & Management*, *43*(6), 756–766. doi:10.1016/j.im.2006.05.001

West, D. M. (2003). *Global e-government, 2003*. Retrieved from http://www.insidepolitics.org/egovt03int.pdf

West, D. M. (2004a). E-government and the transformation of service delivery and citizen attitudes. *Public Administration Review*, *64*(1), 15–27. doi:10.1111/j.1540-6210.2004.00343.x

West, D. M. (2004b). *Global e-government, 2004*. Retrieved from http://www.insidepolitics.org/egovt04int.pdf

West, D. M. (2005). *Global e-government, 2005*. Retrieved from http://www.insidepolitics.org/egovt05int.pdf

West, D. M. (2006). *Global e-government, 2006*. Retrieved from http://www.insidepolitics.org/egovt06int.pdf

West, D. M. (2007). *Global e-government, 2007*. Retrieved from http://www.insidepolitics.org/egovt07int.pdf

Whitmore, A. (2012). A statistical analysis of the construction of the united nations e-government development index. *Government Information Quarterly*, *29*(1), 68–75. doi:10.1016/j.giq.2011.06.003

Witten, I. H., & Frank, E. (2002). Data mining: Practical machine learning tools and techniques with java implementations. *SIGMOD Record*, *31*(1), 76–77. doi:10.1145/507338.507355

World Bank. (2020). *World bank country and lending groups*. Retrieved from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

KEY TERMS AND DEFINITIONS

Beginner Cluster: Beginners are those UN Member States that are novice and in the early stage of e-government development.

Clustering: Clustering is the division of data into groups of similar objects. The goal is to discover the "natural" grouping(s) of a set of patterns, points, or objects.

E-Government Development Index (EGDI): The EGDI is an attempt by the United Nations to assess the extent to which the UN Member States are equipped to deliver government services to clients online and rank them accordingly. The EGDI is a composite measure of three normalized indicators such as provision of online services, telecommunication connectivity, and human capacity.

Electronic Government (E-Government): E-government is the application of IT to government services and procedures to deliver public services and information to citizens, businesses, and governments electronically.

Emergent Cluster: The emergent group is one rank higher than the Beginners and includes those nations that are up-and-coming and growing their e-government service provisions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Information Technology (IT): IT refers to all technologies such as hardware, software, networks, and data storage and management systems that have the capability to process, store, or disseminate data.

Progressor Cluster: This group is one step higher than the Emergent and includes those nations that are progressing and expanding their e-government services.

Unsupervised Machine Learning: The use of machine learning techniques to identify or find hidden patterns or structures without providing output examples. The algorithms discover hidden patterns or data groupings without the need for human intervention. Clustering is one such example.

Veteran Cluster: The Veteran group includes those nations that are proficient and connoisseurs in their e-government development. This group is at the very top in terms of e-government development and service provision.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX 1

Table 5. E-government Profiles of the UN Member States from 2008 to 2020 by Income Group and Cluster Labels

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
1	Andorra	HIGH INCOME	Progressor	Progressor	Progressor	Progressor	Emergent	Progressor	Progressor
2	Antigua and Barbuda		Progressor	Progressor	Veteran	Progressor	Emergent	Emergent	Progressor
3	Australia		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
4	Austria		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
5	Bahamas		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
6	Bahrain		Progressor	Veteran	Progressor	Progressor	Veteran	Veteran	Veteran
7	Barbados		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Progressor
8	Belgium		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
9	Brunei Darussalam		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
10	Canada		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
11	Chile		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
12	Croatia		Progressor	Progressor	Veteran	Progressor	Progressor	Progressor	Progressor
13	Cyprus		Progressor	Progressor	Progressor	Progressor	Emergent	Veteran	Veteran
14	Czech Republic		Veteran	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
15	Denmark		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
16	Estonia		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
17	Finland		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
18	France		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
19	Germany		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
20	Greece		Progressor	Progressor	Veteran	Progressor	Progressor	Progressor	Veteran
21	Hungary		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Progressor
22	Iceland		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
23	Ireland		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
24	Israel		Veteran	Veteran	Veteran	Veteran	Progressor	Veteran	Veteran
25	Italy		Veteran	Veteran	Veteran	Progressor	Veteran	Veteran	Veteran
26	Japan		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
27	Kuwait		Progressor	Progressor	Progressor	Progressor	Veteran	Progressor	Progressor
28	Latvia		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
29	Liechtenstein		Veteran	Veteran	Veteran	Progressor	Veteran	Veteran	Veteran
30	Lithuania		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Veteran
31	Luxembourg		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
32	Malta		Progressor	Progressor	Progressor	Progressor	Veteran	Veteran	Veteran
33	Mauritius		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
34	Netherlands		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
35	New Zealand		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
36	Norway		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
37	Oman		Emergent	Emergent	Progressor	Emergent	Emergent	Progressor	Progressor
38	Panama		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
39	Poland		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
40	Portugal		Progressor	Progressor	Veteran	Progressor	Progressor	Veteran	Veteran
41	Qatar		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor
42	Republic of Korea		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
43	Romania		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
44	Saint Kitts and Nevis		Progressor	Progressor	Progressor	Progressor	Emergent	Progressor	Progressor
45	Saudi Arabia		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
46	Seychelles		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
47	Singapore		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
48	Slovakia		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor
49	Slovenia		Veteran	Veteran	Veteran	Progressor	Progressor	Progressor	Veteran
50	Spain		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
51	Sweden		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
52	Switzerland		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
53	Trinidad and Tobago		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
54	United Arab Emirates		Progressor	Progressor	Progressor	Progressor	Veteran	Veteran	Progressor
55	United Kingdom of Great Britain and Northern Ireland		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
56	United States of America		Veteran	Veteran	Veteran	Veteran	Veteran	Veteran	Veteran
57	Uruguay		Progressor	Progressor	Progressor	Progressor	Progressor	Veteran	Veteran

continued on following page

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Continued

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
58	Albania	UPPER MIDDLE INCOME	Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
59	Argentina		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
60	Armenia		Emergent	Emergent	Progressor	Emergent	Emergent	Emergent	Progressor
61	Azerbaijan		Emergent	Emergent	Progressor	Emergent	Emergent	Progressor	Progressor
62	Belarus		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
63	Belize		Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
64	Bosnia & Herzegovina		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
65	Botswana		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
66	Brazil		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
67	Bulgaria		Progressor	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
68	China		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
69	Colombia		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
70	Costa Rica		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
71	Cuba		Beginner	Emergent	Beginner	Beginner	Beginner	Beginner	Beginner
72	Dominica		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
73	Dominican Republic		Progressor	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
74	Ecuador		Progressor	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
75	Equatorial Guinea		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
76	Fiji		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
77	Gabon		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
78	Georgia-Country		Emergent	Emergent	Progressor	Emergent	Emergent	Progressor	Progressor
79	Grenada		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
80	Guatemala		Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
81	Guyana		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Emergent
82	Indonesia		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
83	Iran-Islamic Republic		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
84	Iraq		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
85	Jamaica		Progressor	Progressor	Emergent	Emergent	Emergent	Emergent	Emergent
86	Jordan		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
87	Kazakhstan		Emergent	Progressor	Progressor	Progressor	Progressor	Progressor	Veteran
88	Lebanon		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
89	Malaysia		Progressor	Progressor	Progressor	Emergent	Emergent	Progressor	Progressor
90	Maldives		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
91	Mexico		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
92	Montenegro		Beginner	Progressor	Progressor	Progressor	Progressor	Progressor	Progressor
93	Namibia		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Emergent
94	North Macedonia		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
95	Paraguay		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
96	Peru		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
97	Russian Federation		Progressor	Progressor	Veteran	Progressor	Progressor	Veteran	Veteran
98	Saint Lucia		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
99	Saint Vincent and the Grenadines		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent
100	Samoa		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner
101	Serbia		Progressor	Progressor	Progressor	Emergent	Progressor	Progressor	Progressor
102	South Africa		Progressor	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
103	Suriname		Emergent	Emergent	Progressor	Emergent	Emergent	Emergent	Emergent
104	Thailand		Progressor	Progressor	Emergent	Emergent	Emergent	Progressor	Progressor
105	Tonga		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
106	Turkey		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
107	Turkmenistan		Beginner	Beginner	Beginner	Emergent	Beginner	Beginner	Emergent
108	Venezuela		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Emergent

continued on following page

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Continued

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
109	Algeria	LOWER MIDDLE INCOME	Emergent	Emergent	Emergent	Beginner	Beginner	Emergent	Emergent
110	Angola		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
111	Bangladesh		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
112	Benin		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
113	Bhutan		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
114	Bolivia		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
115	Cambodia		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
116	Cameroon		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
117	Cape Verde		Emergent	Emergent	Emergent	Emergent	Emergent	Veteran	Veteran
118	Comoros		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
119	Congo		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
120	Coted Ivoire		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
121	Djibouti		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
122	Egypt		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
123	El Salvador		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Emergent
124	Ghana		Beginner	Beginner	Beginner	Beginner	Beginner	Emergent	Emergent
125	Honduras		Emergent	Emergent	Emergent	Emergent	Beginner	Beginner	Beginner
126	India		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
127	Kenya		Beginner	Beginner	Emergent	Beginner	Beginner	Beginner	Emergent
128	Kyrgyzstan		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
129	Lao Peoples Democratic Republic		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
130	Lesotho		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
131	Mauritania		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
132	Mongolia		Emergent	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
133	Morocco		Emergent	Emergent	Emergent	Emergent	Emergent	Beginner	Emergent
134	Nepal		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
135	Nicaragua		Emergent	Emergent	Beginner	Beginner	Beginner	Beginner	Emergent
136	Nigeria		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
137	Pakistan		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
138	Papua New Guinea		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
139	Philippines		Progressor	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
140	Republic of Moldova		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
141	Sao Tome and Principe		Emergent	Emergent	Beginner	Beginner	Beginner	Beginner	Beginner
142	Senegal		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
143	Solomon Islands		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
144	Sri Lanka		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
145	Timor-Leste		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent
146	Tunisia		Emergent	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
147	Ukraine		Progressor	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
148	United Republic of Tanzania		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
149	Uzbekistan		Beginner	Emergent	Emergent	Emergent	Emergent	Emergent	Progressor
150	Vanuatu		Beginner	Beginner	Emergent	Beginner	Beginner	Beginner	Emergent
151	Vietnam		Emergent	Progressor	Progressor	Emergent	Emergent	Emergent	Progressor
152	Zimbabwe		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Emergent

continued on following page

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Continued

#	Country Name	Income Group	2008	2010	2012	2014	2016	2018	2020
153	Afghanistan	LOW INCOME	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
154	Burkina Faso		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
155	Burundi		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
156	Chad		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
157	Democratic Republic of the Congo		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
158	Ethiopia		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
159	Gambia		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
160	Guinea-Bissau		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
161	Haiti		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
162	Liberia		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
163	Madagascar		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
164	Malawi		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
165	Mali		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
166	Mozambique		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
167	Niger		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
168	Rwanda		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
169	Sierra Leone		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
170	Sudan		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
171	Syrian Arab Republic		Emergent	Emergent	Emergent	Beginner	Beginner	Beginner	Emergent
172	Tajikistan		Beginner	Beginner	Emergent	Emergent	Beginner	Beginner	Emergent
173	Togo		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
174	Uganda		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner
175	Yemen		Beginner	Beginner	Beginner	Beginner	Beginner	Beginner	Beginner

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 5
Artificial Intelligence as Driver for SME Competitiveness

Nicola Del Sarto
Scuola Superiore Sant'Anna, Italy

Andrea Piccaluga
Scuola Superiore Sant'Anna, Italy

ABSTRACT

Artificial intelligence is profoundly changing the way in which companies compete and do business. In particular, artificial intelligence can represent a very interesting opportunity for small and medium-sized enterprises, which are constantly looking for new technologies to be able to remain competitive in a turbulent market. However, research exploring how SMEs may successfully adopt artificial intelligence technology are missing. To address this gap, the authors reviewed the literature on artificial intelligence and identified four key features that SMEs need to consider when implementing this technology represented by people, processes, products, and customers. After that they described four cases of Italian SMEs which have adopted successfully this new technology taking into account one of the four dimensions. The study contributes to the literature on artificial intelligence and SMEs and may be helpful for managers who want to adopt this technology within their company.

INTRODUCTION

Artificial intelligence (AI) is a computer technology that revolutionizes the way humans interact with machines, and machines with each other (Yams et al., 2020). Artificial intelligence provides a robot with computing qualities that allow it to perform complex operations and "reasoning", until recently exclusive characteristics of human reasoning, in a short time.

Artificial intelligence is already part of our daily life and is now used in more than half of large companies. Everyone is familiar with driverless cars or voice assistants like Apple's Siri, Microsoft's Cortana, or Google's Alexa, but there are many lesser-known examples. Intelligent algorithms, capable of self-learning, suggest products to buy, films or music tracks in line with our tastes, they know how to

DOI: 10.4018/978-1-7998-6985-6.ch005

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

answer customer questions via chat, they can recognize a person's face to enable access, sort documents based on content, support doctors in reading X-ray images and diagnoses, filter CVs to select the ideal candidate (Davoyan, 2020; Gautier et al., 2020).

There are already numerous examples of how the introduction of artificial intelligence in business processes has led to positive impacts, automating repetitive and low added value parts of the processes themselves, previously carried out by man, reducing errors, allowing the development of new products and services (Wang, & Alexander, 2016). We are at the beginning, but in the next few years, we will see a quantum leap.

The major corporations are taking the first steps in this direction: the first to arrive at concrete results will benefit from a huge competitive advantage. According to the consultancy agencies Accenture and Forrester, 2025 will be the year in which such solutions will explode. Over the past two years, we have seen many companies learn to fail quickly in implementing AI initiatives, but in the future we expect them to seize the opportunities, thanks to more mature approaches.

Not only large companies, digitization is also involving small and medium-sized enterprises which are increasingly oriented towards the adoption of new technologies (Haefner et al., 2020). The introduction of such new technologies may help them overcome both liabilities of newness and adolescence which the literature has identified as crucial problems for small and medium enterprises (Strotmann, 2007). This technology adoption must include at some point, artificial intelligence technologies (Gebauer et al., 2020).

In the coming years, the junction point between the various technologies will be artificial intelligence and machine learning (Li et al., 2020). These technologies range from automation to the constantly connected objects of the Internet of Things (IoT), passing through cloud computing (Atlam et al., 2020). All those technologies will make machines more and more autonomous, meaning they will be able to make decisions without depending on the human factor (Farrokhi et al., 2020; Trunk, Birkel, & Hartmann, 2020). Moreover, they will also be able to interact with the surrounding environment, work on a human factor. Artificial intelligence and machine learning are changing the competitive landscape and several businesses must be ready to face this revolution (Benzidia, Makaoui, & Bentahar, 2020; Makarius et al., 2020).

Not approaching artificial intelligence technologies in the right way may be risky for small and medium enterprises (SMEs) as it can nullify the advantages brought by smart processing techniques. For a small and medium-sized business, this means understanding how AI systems can add value (Urbinati et al., 2019).

Despite previous literature has analyzed the technological aspect and the impact of artificial intelligence on business models and the process of value creation in general (Langley et al., 2020), only a few academic studies have analyzed how artificial intelligence may be beneficial for SMEs and what are the key features that they need to consider to successfully implement such new technology.

We attempt to fill this gap by exploring the following research questions: Does Artificial intelligent help SMEs' competitiveness? What key elements are crucial for successfully implement such new technology?

To address these research questions, we conduct an in-depth analysis of the literature on AI and examine the most relevant elements that companies must consider when implementing AI. In doing so, we propose a table that points to the importance of four key dimensions: 1) People; 2) Processes; 3) Products; 4) Customers. After, we try to validate this matrix empirically, considering four representative cases of SMEs that successfully implemented AI technologies by considering the four elements identified by the literature. Those companies were all situated in Italy, and in particular in the northeastern, central, southern, and north-west. We decided to focus on Italy as this is considered a country in which

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the manufacturing sector represents the core activity of the production system (Goodman et al., 2016) which will be strongly influenced by these emerging technologies (Caputo et al., 2016).

We find that to succeed in the process of AI technology adoption, SMEs must focus on one of the dimensions identified and exploit it accordingly to the characteristics of the business of the clients and of the firm itself. We believe that such results may be beneficial for SMEs which are willing to adopt this emerging technology. In particular, our study can be helpful for SME managers who may be inspired by the success stories of analyzed firms.

Following this introduction, we structured the article as such. First, we provide the theoretical background in which we reviewed relevant literature on AI. Then we illustrated the methodology and we discuss our finding. In the conclusion section, we summarize the contributions, limitations, and future research of the paper.

BACKGROUND

The Impact of Artificial Intelligence on Business

The fourth industrial revolution confronts us with a digital transformation process that aims to equip the company with an organizational and process structure capable of rapidly and continuously introducing new technologies to support the innovation of their processes (Canhoto, & Clear, 2020; Fernández-Rovira, et al., Forthcoming; Wang et al., 2020). Production, assembly, and distribution systems have always produced enormous quantities of data, and today, thanks to intelligent processing and analysis systems, they can be used productively, along the entire value chain and assuming an increasingly decisive role of guide in the decision-making systems of the company (Angelis, & da Silva, 2019; Farrokhi et al., 2020).

The introduction of AI systems into any business process can be comparable to the impact that the introduction of mechanization has had in industries since the mid-eighteenth century (Garbuio, & Lin, 2019). A decisive step in many production sectors that improved the conditions of workers engaged at the time in repetitive or strenuous work, replacing man or animal strength with the mechanical one, generated by combustion from a motor machine, originating a whole series of changes, economic and social (Benzidia, Makaoui, & Bentahar, 2020). Artificial Intelligence, therefore, has the objective, or rather ambition, to emulate the cognitive abilities of human beings, increasing their effectiveness (Hwang, 2016). Let's think, for example, of forecasting systems that aim to predict the production demand of a plant by analyzing not only what is produced over a given period through projections, but also the influence that unpredictable external variables, such as oscillation of the market or competition, have on the generation of demand (Haefner et al., 2020).

To date, there are several noteworthy use cases in many areas and industries such as in the healthcare (Intelligent Health), financial (Financial Trading), automotive (Autonomous Vehicle, Connected Car) sectors, as well as scenarios that are now in use common as security and emergency (Image Processing, Computer Vision, Object Detection and Recognition) and also concerning marketing customizations (Natural Language Processing for Sentiment Analysis) (Bogachov et al., 2020). AI intelligence allows robots to perform complex jobs with very high efficiency, but these automated objects are often made to perform a single task and the reprogramming cost is very high, if possible (Davis et l., 2020). We are talking about automation, not AI, but the first adaptive collaborative robots are already being thought and designed, equipped with an AI capable of learning different tasks through learning by demonstra-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion and hardware more suitable for the reconfigurability of activities (Wang, Luo, & Huang, 2020). Adaptive manufacturing is also strategic for the sector and requires AI solutions capable of adapting to the scenario and flexible hardware infrastructures capable of interacting with people in the same environment (Richardson, 2020).

In production plants, we refer more often to Machine Learning for the management of anomalies and predictive maintenance (Di Vaio, et al., 2020). The next step is to have production systems capable of intervening autonomously based on experience and that do not limit themselves to intervening on the system when thresholds are exceeded, and according to pre-established rules, but which learn from the analyzes previously carried out and create the representation of the production process according to "unscheduled" variables. This is the kind of intelligence necessary for the interaction of complex integrated automatic systems, for example, to ensure the operation of a highly automated extended supply chain and is one of the main investment areas of many companies. AI brings tangible and quantifiable benefits such as, for example, support to the workforce (skills rotation), process automation (working capital), customer management (customer loyalty), and product innovation (product servitization) (Tanev, & Sandstrom, 2019).

It is important to report that over the next few years, clarifying and regulating the mission and the areas of use, especially from an ethical point of view, will be fundamental. Trust in AI will be earned over time as it happens with human relationships (Wang, Luo, & Huang, 2020). It will be necessary to demonstrate that it is not a "human against machines" struggle by going beyond fears and clarifying the doubts related to adoption thanks to the results achieved. AI can complement the work of human personnel, increase judgment and analysis capacity and skills, allowing them to focus resources to significantly expand ingenuity, creative effort, and experience, improving their speed, scope, and efficiency.

By analyzing literature in the academic database Scopus and using the research query ("Artificial Intelligence" OR "AI" OR "Machine Learning") in the period 2010/2020, we identified relevant literature. Management scholars have identified six main areas in which Artificial intelligence may be disentangled. Scholars pointed out how these 6 areas are intertwined with emerging technology.

- Automation: in this context, this term refers to any solution that allows people to be relieved of repetitive and complex actions, but which lend themselves to being structured and implemented by a computer or mechanical system. We often speak of "industrial automation" referring to technologies that allow the automation of part or entire production processes, but the term automation can also be associated with contexts other than industrial. Consider, for example, the delivery processes of services, logistics, and the tracking of goods. The introduction of AI technologies opens up new frontiers in the field of automation allowing you to face processes with a complex decision-making component.
- Computerization: the introduction of software applications and the necessary infrastructures, computers, servers, and networks within a company, to automate or make processes more efficient. Computerization is a basic requirement to be able to exploit the potential of AI, if only because it enables the possibility of producing and collecting structured and unstructured data from which AI can contribute to extract useful information for the improvement of processes or the definition of new business models.
- Dematerialization: the replacement of paper documents with digital ones, not only as a consequence of the computerization process but as a driver for the review of business processes to make them more efficient, more effective, safer, and sustainable. The tools for tracing and protecting the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

rights of access to digital documents, also as the new general data protection regulation (GDPR) regulation, today represent an essential tool.

- Virtualization: we could define it in general as the set of technologies that allow you to make the most of the processing capabilities of a hardware system. More technically, the ability to abstract hardware resources (CPU, RAM, and Storage) making them available to the software in a virtual way. Wanting to give a very simple example, with virtualization technologies it is possible to simulate the execution of multiple virtual servers on a single physical server that can work simultaneously by exploiting the power of the hardware more efficiently.
- Cloud Computing & Big Data: Cloud Computing means the possibility of being able to use hardware and software resources in a "service" logic by paying, for example, only for the time of actual use. The potential of cloud computing goes beyond mere economic efficiency, it is possible to face the need for a sudden increase in processing capacity, to easily implement mechanisms to increase the availability of services (Disaster Recovery and Business Continuity), and to have adequate security requirements guaranteed. to the reference application context. By Big Data we mean the set of technologies and methodologies capable of processing huge amounts of structured and unstructured data and extracting useful information for corporate decision-making processes.
- Mobile: one of the most disruptive changes in recent years is certainly the spread of mobile devices both in everyday life and in business life. The evolution of wireless telecommunications networks and the now imminent introduction of 5g in mobile telephony represent one of the most important factors in the digital transformation processes of companies. The impact in the business world is extended to all areas, just think of smart working and the IoT (Atlam et al., 2020). AI is already a reality in this world, think of the virtual assistants that are now on all smartphones and of the device access protection techniques based on facial recognition.

ARTIFICIAL INTELLIGENCE AND SMEs

Not just big companies are dealing with digitalization, this process is also involving small and medium-sized enterprises, which are increasingly oriented towards the technology of Industry 4.0. The focus on these firms is justified by the fact that SMEs are characterized by peculiar features. Moeuf et al. (2020) identified four managerial characteristics of SMEs: short-term strategy, the importance of the SME manager, lack of expert support function, and short hierarchical line. Moreover, Laforet and Tann (2006) found four key characteristics: culture, process, leadership, and company strategy, as the important factors. Notwithstanding the adoption of new technologies is crucial for the achievement of superior competitive advantage (Pappas et al., 2021) SMEs remain cautious about adopting the solutions offered by digital technologies (Crupi et al., 2020). According to Sommer (2015), there is an inverse relationship between the dimension of SMEs and the impact that the digital revolution may have on them. Radziwon et al. (2014) However, at the moment investments and development of capabilities in automation, are not developed enough by SMEs, moreover, lack of resources and perceived uncertainty act as key obstacles for SMEs to participate in digital manufacturing at the same rate as larger companies. Despite these obstacles, AI is likely to become a complementary tool for decision-making in SMEs (Magistretti et al. 2019) and for this reason, its adoption is recognized as crucial (Häfner et al., 2020). From automation to the always-connected objects of the Internet of Things, passing through cloud computing: in the coming

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

years, the junction point between the various technologies will be artificial intelligence and machine learning (Sestino et al., 2020).

Machines will become more and more autonomous, they will be able to make decisions without depending on the human factor (Trunk, Birkel, & Hartmann, 2020). They will also be able to interact with the surrounding environment, work alongside individuals, learn, and above all, they will be able to manage and control production processes. Artificial intelligence and machine learning factories are changing and businesses will have to be ready to face this revolution (Obschonka, & Audretsch, 2019).

The data collected by intelligent machines is one of the most important elements (McAfee et al., 2012). They will allow companies not only to keep production processes under control and to make forecasts, thanks to Analytics, but will also be crucial to perfecting devices designed for artificial intelligence. The latter is equipped with neural networks - comparable to the human brain - which thanks to Big Data they learn and improve (Lee, 2018). It is essential for the company to correctly indicate to the machines where and how this information is stored. Switching to intelligent processing techniques also requires investing in personnel. A part will necessarily have to be redeveloped so that it knows how to move in the smart factory. Then, small and medium-sized enterprises must consider the possibility of hiring new figures, specially designed to guide the change of the company and to manage new working methods.

The massive change in the speed of exchanging information, as well as the volume, variety, and voracity with which data guide immediate access to technology in our daily lives, represent the new characteristics that will distinguish Italian know-how in the future. Over the years, we have been accustomed to depicting our company as a horizontal structure, within which all phases of value creation follow one another linearly and distinctly (Gans, & Ryall, 2017). Generally speaking, we start from the research and development, production, logistics, distribution, up to the sales and customer service. Thanks to new technologies and mainly data, this business model will no longer have any value in the short term: it will be considered obsolete. AI allows the company to be modeled as a (circular) platform, in which data, available in real-time, is at the center of the organizational model (Di Vaio, et al., 2020).

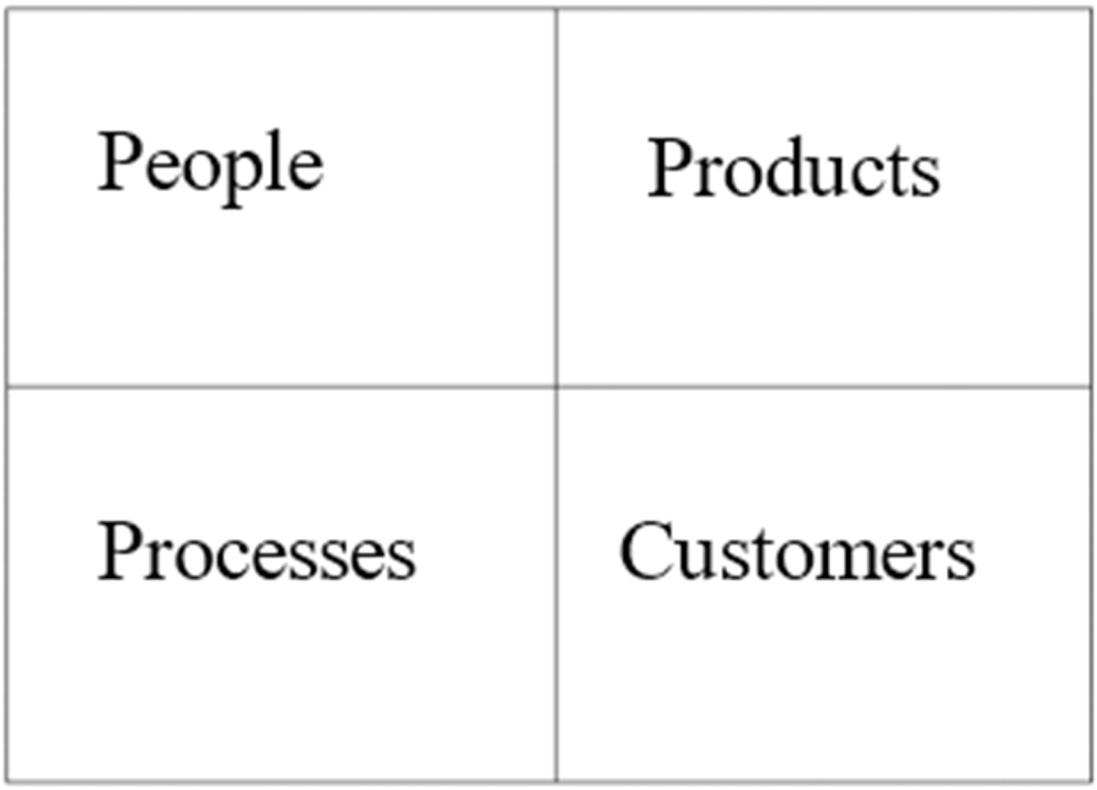
Data is the key to understanding this new business organization (Davenport et al., 2012), which leverages four distinct but strongly interconnected areas thanks to AI illustrated in table 1. The first two areas are people and processes, these are areas within the company, they guide actions and aim to create efficiency (Chang, 2020). Next, we have customers and products, external areas that have the task of creating effectiveness in entrepreneurial action. To make this new idea of an organizational model based on data concrete and measurable, the following are some use cases of companies that have already embarked on the digital transformation process and the benefits they have achieved over time (Fernández-Rovira, et al., Forthcoming; Shrestha, Krishna, & von Krogh, 2020).

When dealing with human resources, the challenge that SMEs face daily is the management of skills. There are very few companies that today can afford to act on corporate roles, rather than on individuals. We are used to investing in the strong specialization of individuals (How-To), rather than democratizing knowledge (Know-How) horizontally, using for example AI to equip the workforce with tools that mitigate the turn-over during the course.

When we talk about processes, we refer to a sequence of events that determine the execution and completion of specific actions. In the event of an abnormal situation or error, we are used to analyzing the event when it has already happened (understanding what happened, when, and why), rather than continuously analyzing the entire sequence and preventing error situations thanks to predictive analysis techniques (understand what can happen and manage the problem in advance). If we take this dynamic

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1.



to a corporate scale, it is easy to see the benefits in terms of risk management, volatility, and productivity that we can strive for thanks to what is made available today by AI.

As pointed out by previous literature, to achieve a sustainable competitive advantage, SMEs need to be deeply rooted in the territory, capillary, and focused on the relationship which can create with customers (Shrestha, Krishna, & von Krogh, 2020). This is a particularity that, over the years, has allowed SMEs to bridge the technological and infrastructural gaps in favor of a continuous and lasting bond based on trust. The great challenge faced today is twofold: a) to continue to do well where we have an established position and; b) to win over new customers and global markets. All this, at the company level, translates into the ability to be perceived by our customers not as preferred sellers or suppliers (perception based on costs and opportunities), but as active partners in the success of customers (perception based on performance and advantageous relationships).

Product innovation is fundamental for SMEs, undoubtedly this characteristic distinguishes the way they do business. From a globalization perspective, it is also true that SMEs are not always able to break through the market thanks to excellence. Globalization teaches that quantity often prevails over quality. It does not mean that SMEs need to change how innovate, but that they can make the most of new technologies and AI, not only to make the products smarter, but to make them even more unique and attractive (Shrestha, Krishna, & von Krogh, 2020). Machines will become more and more autonomous, they will be able to make decisions without depending on the human factor (Trunk, Birkel, & Hartmann,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

2020). They will also be able to interact with the surrounding environment, work alongside individuals, learn, and above all, they will be able to manage and control production processes. Artificial intelligence and machine learning factories are changing and businesses will have to be ready to face this revolution (Obschonka, & Audretsch, 2019).

The data collected by intelligent machines is one of the most important elements (McAfee et al., 2012). They will allow companies not only to keep production processes under control and to make forecasts, thanks to Analytics, but will also be crucial to perfecting devices designed for artificial intelligence. The latter is equipped with neural networks - comparable to the human brain - which thanks to Big Data they learn and improve (Lee, 2018). It is essential for the company to correctly indicate to the machines where and how this information is stored. Switching to intelligent processing techniques also requires investing in personnel. A part will necessarily have to be redeveloped so that it knows how to move in the smart factory. Then, small and medium-sized enterprises must consider the possibility of hiring new figures, specially designed to guide the change of the company and to manage new working methods.

The massive change in the speed of exchanging information, as well as the volume, variety, and voracity with which data guide immediate access to technology in our daily lives, represent the new characteristics that will distinguish Italian know-how in the future. Over the years, we have been accustomed to depicting our company as a horizontal structure, within which all phases of value creation follow one another linearly and distinctly (Gans, & Ryall, 2017). Generally speaking, we start from the research and development, production, logistics, distribution, up to the sales and customer service. Thanks to new technologies and mainly data, this business model will no longer have any value in the short term: it will be considered obsolete. AI allows the company to be modeled as a (circular) platform, in which data, available in real-time, is at the center of the organizational model (Di Vaio, et al., 2020).

Data is the key to understanding this new business organization (Davenport et al., 2012), which leverages four distinct but strongly interconnected areas thanks to AI illustrated in table 1. The first two areas are people and processes, these are areas within the company, they guide actions and aim to create efficiency (Chang, 2020). Next, we have customers and products, external areas that have the task of creating effectiveness in entrepreneurial action. To make this new idea of an organizational model based on data concrete and measurable, the following are some use cases of companies that have already embarked on the digital transformation process and the benefits they have achieved over time (Fernández-Rovira, et al., Forthcoming; Shrestha, Krishna, & von Krogh, 2020).

METHODOLOGY

Our theoretical development is complemented with examples of four SMEs presenting peculiar characteristics as they are span across Italy and are recognized as “innovative small-medium enterprises”, moreover they have successfully adopted Artificial Intelligence to increase their competitive integrating the solution into their existing business. We chose these SMEs because they seem to explain each of the dimensions we identified in our theoretical section. Thus, following previous studies (De Massis et al., 2016; Siggelkow, 2007), these “exceptional” cases seem to be a proper choice for discussing and analyzing the phenomenon investigated in our study. The goal of this study, indeed, is to use such illustrative examples to clarify theoretical arguments and relationships showing how conceptual reasoning reported in the interpretative framework is applied (Siggelkow, 2007) instead of relating to an inductive case study.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In developing the cases we drew on multiple sources of information. In particular, we collected data from companies' websites and secondary data sources, such as reports, newspapers, specialized websites in the second half of 2020. To analyze data we used an iterative process, which allowed us to move from data to theory and vice versa (Strauss & Corbin, 1998), and enabled us to refine the framework, illustrating how theoretical concepts work in practice.

RESULTS

Through the analysis of the four cases, we can propose examples of how SMEs should adopt AI to achieve a competitive advantage over other firms. Thanks to AI it is possible (at least this is the ultimate goal) to make machines able to perform complex actions and "reasoning", learn from mistakes, and perform functions that until now were exclusive to human intelligence. Today artificial intelligence is used in the company and beyond, to perform tasks that would require a lot of time for humans. AI is a section of computer and computer linguistics that deals with the formalization of problems and tasks similar to those performed by a person. Nowadays it is a sub-discipline of computer science that deals with studying the theory, techniques, and methodologies that allow you to design both hardware and software systems capable of processing electrical performances that simulate the relevance of human intelligence. The result of the work of AI must not be difficult to distinguish from that carried out by a human with specific skills.

It is now well established that the excellence on which SMEs in the manufacturing sector were aiming is today severely tested by the strong dynamism of the global economy. Over 20 years, we have gone from an economy driven by supply (production for the warehouse) to an economy driven by demand (consumption in real-time). Everyday companies fight against the two forces that characterize the phenomenon of globalization: on the one hand, the competition based on the cost of labor and low quality of emerging economies, on the other the pressure of a global value chain that sees SMEs amid an unequal relationship between multinational customers and suppliers. The particularities of SMEs' know-how, namely Quality, Personalization, and Timing, are no longer sufficient to guarantee a leading position in this new scenario (Bag et al., Forthcoming).

In attempting to introduce these new technologies, SMEs must consider the four elements identified above. In the following, we report four cases of SMEs which have successfully implemented AI considering the four elements.

PEOPLE, THE WORKFORCE

The first company is a company situated in northeastern Italy, specialized in the production and sale of flexible machines and systems for sheet metal working, which thanks to the AI revolutionizes how the service was done. Previous to the introduction of the new technology, once the machine produced was delivered to the customer, after the first test there was no longer any trace of the real state of use and maintenance carried out over time. The configuration and ordinary management activities were not tracked by the customer and, in the case of a service request, the service operators could not do anything other than connect remotely to the system and reactively analyze the event log: often this intervention was not finalized due to the lack of evidence on the actual operation of the machine. In the worst case,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

when the problem was complex, the company had to send a team of specialists on-site and often wait days to get permission to visit some of the North African countries where the company had strong growth of revenue. This situation was considered unsustainable over time and exposed the company to a strong risk of competition with multinational companies in the sector, which had capillary centers of expertise and a network of global partners.

The company has estimated, for each day of inactivity, an average loss per machine of 20,000 euros of non-production for the customer. This situation exposed the entire company to a huge risk of customer loss. The company has fully exploited the potential of AI through two distinct approaches. The first at the service level, by creating a remote diagnostics center which, thanks to the data received in real-time from the machine, compares the real use by the customer with the expected use from the first tested configuration. The result is to proactively calculate, through machine learning techniques, the percentage of the risk of extraordinary assistance, then manage the planning of the activities of the service, acting on a knowledge base not of single individuals but explicit and horizontal. The second investment concerns the provision, for each machine, of a mixed reality solution capable of allowing the customer to establish a remote connection in case of urgency and translate the local language into Italian in real-time and vice versa, using natural language processing techniques (Gautier et al., 2020). The result is to guide the customer himself in solving the problem, avoiding sending teams of technicians, limiting machine downtime, and retain the customer's expertise on the product.

The result is not only economic, the company uses this technology today to train new employees on previous cases, managed through technological innovation with excellent results on the quality of customer service and with a positive impact on turnover. This training activity has allowed the company to develop strong competencies and capabilities focused on customers' needs. By training new employees, the company can now rely on a strong base made of high qualified people which constitutes the basis of its competitive advantage.

THE PROCESSES, THE BACKBONE

In the second case, SMEs of Central Italy specialized in the production and molding of plastic and thermosetting materials, implemented AI technologies to makes planning its winning weapon.

Previously to the introduction of AI, the company's working capital was continually put at risk by the strong exposure in the procurement, production, and warehouse management phase. The company purchases raw materials from multinational companies that manage production campaigns without adequate planning and, at the same time, must ensure a continuous supply of different and increasingly personalized products, according to long-term commercial agreements with end customers.

The company was heavily unbalanced in terms of financial exposure due to sequential and fragmented planning. The different supply chain metrics, such as Inventory Turn Over (ITO), One Time Delivery (OTD), as well as the Overall Equipment Effectiveness (OEE) and Part Per Million (PPM) production metrics were out of control and, worse, the various data collected were discordant and not updated on the entire supply chain. Making decisions in this way was difficult and the company preferred to stock, continuing to produce even when demand slowed with the hope of being ready at the time of a resumption of orders. All this impacted the immobilization of working capital of about 5 million euros each year. The solution was to change the way the processes were organized, thus moving from sequential to competitive planning.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

This technique leverages AI to constantly monitor the performance of various processes, measure and collaborate in real-time, allowing immediate adjustments to planning between vertical elements (machinery, assets, operations) and horizontal elements (work orders, raw materials), shortening response times through the information symmetry typical of a compartmentalized production chain. Ultimately, this allows for quicker and more informed strategic decisions.

The company, therefore, leveraged three process innovations that make massive use of AI: a) IoT; b) demand forecasting; c) intelligent process automation (Atlam et al., 2020). Connecting the production plant through IoT techniques to the different levels of automation has allowed us to have real-time visibility of the production capacity, leveraging on availability, quality, and performance, capable of accurately measuring capacity productivity. Acting therefore on the connection of the plant data, with the data deriving from the management and warehouse systems through demand forecasting techniques, which exploit big data and machine learning algorithms, has allowed having a single data collection and analysis center which, leveraging on historical performances and forecasts, allow to adjust the supply of raw materials and warehouse management, planning the company's production system to the advantage. Finally, the automation of the activities of the entire supply chain through the intelligent process automation has allowed the implementation of virtual agents able to support the work of different people by anticipating the tasks to be completed based on the analysis of daily work: therefore able to suggest the best actions to take based on a specific position of the supplier or customer.

The company is measuring the benefit of this real evolution, positively attacking the working capital necessary to maintain the exact (not the minimum) level of production and inventory according to a production plan calculated over the medium-long term. A 50% reduction in current spending means bringing a benefit of about 2.5ml each year. These costs can translate into a continuous investment in new technology and innovation.

CUSTOMERS, THE ALLIES

The third studied company is a company from southern Italy, the world leader in the construction of machines for the production and packaging of sanitary products which introduced AI solutions to offer to its customers' solutions to problems and not just technologies Before the introduction of AI the sales process of a machine was linear over time and based on a presale, design, and opening of a new order. All managed promptly from the first contact, without however having a sales strategy with added value over time. All company divisions aimed at closing the order, from design to testing, the entire production process had the aim of closing the order, obtaining the signature for testing, and issuing an invoice.

The opportunity, therefore, was to enhance the installed park with a digital sales channel that would give the possibility not to visit the customer every time he showed up for specific needs, but to create a continuous dialogue, on a digital channel (always available and intelligent) which would suggest to the customer which package, spare part or action was appropriate to keep the productivity level of the machine at 100%.

A new agreement with the customer, therefore, is not only based on the product but on the ability of the product (or machine) to actively contribute to the customer's success. In this case, technological innovation played a fundamental role. The company has acted on two parallel fronts: having a digital platform connected to the customer's production system and introducing AI within the sales force. The platform relies on a big data database that can receive not only the data of the machine connected via the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

IoT but the production data of the customer himself, thus creating a very intimate level of communication with the customers' production system. The platform is, therefore, able to measure, through machine learning techniques, the impact of the machine within the production system (availability, performance, and quality of the parts produced).

With this platform, the sales force can have valuable information on what the customer needs to further improve performance (additional packages, corrective actions, software services), or anticipate scheduled maintenance (predictive maintenance, anomaly analysis). This information is sent back to the customer as smart alerts that advise the customer to evaluate these options and also have offers on spare parts in real-time.

The salesforce, as well as the customer service, can always be informed about the current state of the machine performance and informed based on all interventions (sales opportunities or repairs) for each customer. The company has estimated the increase in customer continuity at + 30%, which shows that the company has focused on the sale of value to the customer (continuous and predictable cycle) rather than focusing on technical presales (occasional unpredictable cycle).

THE PRODUCTS, THE EXCELLENCE

The importance of the excellence of the product is well recognized, particularly for SMEs. The last studied company is a company from north-west Italy, a world leader in the manufacture of electrical components and switches, which transforms thanks to AI a commodity into a strategic asset.

The company produces products that are present in the daily life of businesses and buildings and for years has had an important market share also on an international scale. The product catalog was very large, with thousands of very similar products that differed only in the voltage load and other small differences in terms of product mechanics. Customers, mostly wholesale channels and installation companies, chose the product not only for its technical equipment, but also for the ease with which they were able to find the product on the market and, therefore, the speed with which they could benefit from the product for their commercial purposes. This type of market pushed the company to release more and more products of a similar type (thousands of versions in the catalog) and to have a wholesale pricing strategy based on sales volumes for certain categories of customers.

AI made it possible to play a new game on a new table, where precisely the products would become real strategic assets for customers. First, the company consolidated the number of products into a few dozen, bridging the mechanical and electrical characteristics with intelligent software solutions that were installed on the products but enabled only based on commercial negotiations. This approach has meant that the role of product performance enabler (resistance, block, energy saving), previously imaginable only through the use of hardware or mechanical components, has been transferred to AI.

Secondly, the company has created a series of intelligent software solutions that it has offered to its customers for free (product configuration using mixed reality, electrical cabinet configuration using virtual assistants, projection of plant revenues using machine learning techniques), to divert the customer's attention to the mere price of the product and focusing attention on the company's ability to speed up the work of its customers effectively.

The introduction of this new solution based on artificial intelligence allowed the company to transform its product from commodity to a strategic asset, thus improving the relationship with the customer and offering him a more complete service than that previously offered.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

DISCUSSION

AI is a set of multiple technologies that allow machines to perceive, understand, act and learn, both independently and to enhance human activities. Important data and studies reveal a daunting truth about today's growth (Bag et al., Forthcoming). To push economic growth, there has been a sharp reduction in the traditional levers of production, capital, and labor investments. However, it is not just a matter of data and studies. AI is a new factor of production and has the potential to introduce new sources of growth, changing the way work is perceived and strengthening the role of people driving job growth.

Research on the impact of AI in 12 major economic sectors reveals that AI could double annual economic growth rates in 2035, thus changing the nature of work and creating a new relationship between man and machine. The impact of AI technologies on work is estimated to lead to an increase in work productivity by up to 40% and enable people to make more efficient use of their time.

This impact is particularly important for SMEs, that must be able to leverage such emerging technologies in a more dynamic and globally connected competitive landscape. According to numerous scholars, the correct use of these technologies will allow SMEs to achieve a sustainable competitive advantage (Baldegger, Caon, & Sadiku, 2020). However, research on how SMEs may successfully adopt AI technologies are missing.

For this reason, we try to fill this gap by illustrating examples of four SMEs that have successfully adopted AI considering the crucial dimensions pointed out by the literature (i.e. People, Processes, Products, Customers). We think our study may contribute to the literature investigating the importance of AI for SMEs, illustrating four successful cases of SMEs which adopted AI technologies and integrated them into their business. In particular, such companies have successfully managed with these technologies by considering four key aspects represented by People, Processes, Products, and Customers. In the first case, the firm has adopted AI to encourage staff training by making them more competent in managing customer care, thus forming distinctive skills which now represents a strong asset for the SME. In the second case, AI technologies allowed the transformation of internal processes, making them more efficient and flexible to the demands of a constantly evolving market. In the third case, AI has been introduced to deliver innovative products by adding new features and services able to increase customers' experience and satisfaction. The fourth case represents the case of an SME that introduced AI to strengthen its connection with customers, thus increasing customer retention and enlarging the customer base.

The context in which all companies operate is transforming due to the growing competitiveness and turbulence of today's markets and that is why manufacturing companies, traditionally structured with product-focused business models, now tend to rely on models that are increasingly oriented towards service in which the customer is placed at the center to get to supply products that are best suited to their specific needs. As also pointed out previously, the services business allows to establish lasting relationships with customers and allows to create the ideal conditions for establishing long-term relationships of mutual trust that stimulate loyalty and increase the propensity to buy back, bringing an advantage to company difficult to understand and copy by the competition.

The Customer Experience is therefore a fundamental element for building advantageous relationships with customers and involves the centrality of the customer in all phases of the relationship with companies, before and after the sale. The conversation with the customer must be continuous and effective, all the channels now widely disseminated must be available, and also all the information possessed about the customer: this opening exposes companies to continuous solicitation and evaluation from the outside.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Furthermore, the contacts developed during the production process between companies, customers, and suppliers end up constituting a mine of information that must be capitalized.

The Industry 4.0 initiatives have disseminated, to a wide audience, the knowledge that there are new technologies available today, necessary to build new industrial solutions, highlighting how the concepts of integrated design, connectivity, and interoperability are the indispensable complement to achieve effective innovation, which guarantees national and extra-national competitiveness and allows, not only the transformation of the business model but also the emergence of new models (Davis et al., 2020).

The awareness of the existence of these technologies is not in itself understanding them and how they must be used synergistically to obtain a result. The need to convey a strong message to SME entrepreneurs, accompanied by fiscal support to start projects, has focused on the technological aspects; next years will be fundamental for the adoption of AI services and related business models, and for creating a scalable model, in which the market recognizes the value of these new technologies and uses them in the most suitable forms.

CONCLUSIONS

Automation with inherent AI is a growing field with diverse applications in various industries. Our study contributes to a richer understanding of how SMEs can successfully apply artificial intelligence in their traditional business to achieve competitive advantage. To fill this gap in this research we answer the following research questions: "Does Artificial intelligent help SMEs competitiveness? What key elements are crucial for successfully implement such new technology?. Through the analysis of four cases of Italian SMEs which adopted AI, we shed light on the four main dimensions which need to be considered to successfully adopt the new technology and we deeply described how these SMEs have done. We believe that our research provides both theoretical and managerial implications. We contribute to the literature on SMEs by showing how these particular firms may use new technologies to overcome their liabilities (Strotmann, 2007) and achieve sustainable competitive advantage. For what concerns managerial contributions, we present four cases that may serve as inspirations for managers who want to adopt artificial intelligence in their own SME.

As with all research, this study has limitations that provide avenues for future research in the growing field of research on SMEs and Artificial Intelligence. In particular, future research must try to validate our framework in another country, thus validating it, generalizing the results, and comparing them across the country. Moreover, scholars should consider the adoption of some quantitative methods, collecting data through ad hoc surveys and secondary data thus enlarging and deepening the research field.

REFERENCES

Angelis, J., & da Silva, E. R. (2019). Blockchain adoption: A value driver perspective. *Business Horizons*, *62*(3), 307–314. doi:10.1016/j.bushor.2018.12.001

Atlam, H. F., Azad, M. A., Alzahrani, A. G., & Wills, G. (2020). A Review of Blockchain in Internet of Things and AI. *Big Data and Cognitive Computing*, *4*(4), 28. doi:10.3390/bdcc4040028

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bag, S., Pretorius, J. H. C., Gupta, S., & Dwivedi, Y. K. (forthcoming). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technological Forecasting and Social Change, 163*, 120420.

Baldegger, R., Caon, M., & Sadiku, K. (2020). Correlation between Entrepreneurial Orientation and Implementation of AI in Human Resources Management. *Technology Innovation Management Review*, *10*(4), 72–79. doi:10.22215/timreview/1348

Benzidia, S., Makaoui, N., & Bentahar, O. (2020). The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. *Technological Forecasting and Social Change*, *165*, 120557. doi:10.1016/j.techfore.2020.120557

Bogachov, S., Kwilinski, A., Miethlich, B., Bartosova, V., & Gurnak, A. (2020). Artificial intelligence components and fuzzy regulators in entrepreneurship development. *Entrepreneurship and Sustainability Issues*, *8*(2), 487–499. doi:10.9770/jesi.2020.8.2(29)

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons*, *63*(2), 183–193. doi:10.1016/j.bushor.2019.11.003

Caputo, A., Marzi, G., & Pellegrini, M. (2016). The Internet of Things in manufacturing innovation processes: Development and application of a conceptual framework. *Business Process Management Journal*, *22*(2), 383–402. doi:10.1108/BPMJ-05-2015-0072

Chang, K. (2020). Artificial intelligence in personnel management: The development of APM model. *The Bottom Line (New York, N.Y.)*, *33*(4), 377–388. doi:10.1108/BL-08-2020-0055

Crupi, A., Del Sarto, N., Di Minin, A., Gregori, G. L., Lepore, D., Marinelli, L., & Spigarelli, F. (2020). The digital transformation of SMEs–a new knowledge broker called the digital innovation hub. *Journal of Knowledge Management*, *24*(6), 1263–1288. doi:10.1108/JKM-11-2019-0623

Davenport, T. H., Barth, P., & Bean, R. (2012). How Big Data Is Different. *MIT Sloan Management Review*, *54*(1), 43.

Davis, R., Vochozka, M., Vrbka, J., & Neguriţă, O. (2020). Industrial Artificial Intelligence, Smart Connected Sensors, and Big Data-driven Decision-Making Processes in Internet of Things-based Real-Time Production Logistics. *Economics. Management and Financial Markets*, *15*(3), 9–15.

Davoyan, A. (2020). The Impact of Artificial Intelligence on Work, Education, Mobility and Economy. In *Proceedings of the Future Technologies Conference* (pp. 291-296). Springer.

De Massis, A., Frattini, F., Kotlar, J., Petruzzelli, A. M., & Wright, M. (2016). Innovation through tradition: Lessons from innovative family businesses and directions for future research. *The Academy of Management Perspectives*, *30*(1), 93–116. doi:10.5465/amp.2015.0017

Di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. (2020). Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, *121*, 283–314. doi:10.1016/j.jbusres.2020.08.019

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Farrokhi, A., Shirazi, F., Hajli, N., & Tajvidi, M. (2020). Using artificial intelligence to detect crisis related to events: Decision making in B2B by artificial intelligence. *Industrial Marketing Management*, *91*, 257–273. doi:10.1016/j.indmarman.2020.09.015

Fernández-Rovira, C., Valdés, J. Á., Molleví, G., & Nicolas-Sans, R. (forthcoming). The digital transformation of business. Towards the datafication of the relationship with customers. *Technological Forecasting and Social Change, 162*, 120339.

Gans, J., & Ryall, M. D. (2017). Value capture theory: A strategic management review. *Strategic Management Journal*, *38*(1), 17–41. doi:10.1002mj.2592

Garbuio, M., & Lin, N. (2019). Artificial intelligence as a growth engine for health care startups: Emerging business models. *California Management Review*, *61*(2), 59–83. doi:10.1177/0008125618811931

Gautier, A., Ittoo, A., & Van Cleynenbreugel, P. (2020). AI algorithms, price discrimination and collusion: A technological, economic and legal perspective. *European Journal of Law and Economics*, *50*(3), 405–435. doi:10.100710657-020-09662-6

Gebauer, H., Arzt, A., Kohtamäki, M., Lamprecht, C., Parida, V., Witell, L., & Wortmann, F. (2020). How to convert digital offerings into revenue enhancement–Conceptualizing business model dynamics through explorative case studies. *Industrial Marketing Management*, *91*, 429–441. doi:10.1016/j.indmarman.2020.10.006

Goodman, E., Bamford, J., & Saynor, P. (Eds.). (2016). *Small firms and industrial districts in Italy*. Routledge.

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2020). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, *162*, 120392. doi:10.1016/j.techfore.2020.120392

Hwang, J. S. (2016). The fourth industrial revolution (industry 4.0): Intelligent manufacturing. *SMT Magazine*, *3*, 616–630.

Laforet, S., & Tann, J. (2006). Innovative characteristics of small manufacturing firms. *Journal of Small Business and Enterprise Development*, *13*(3), 363–380. doi:10.1108/14626000610680253

Langley, D. J., van Doorn, J., Ng, I. C., Stieglitz, S., Lazovik, A., & Boonstra, A. (2021). The Internet of Everything: *Smart things and their impact on business models. Journal of Business Research*, *122*, 853–863. doi:10.1016/j.jbusres.2019.12.035

Lee, H. L. (2018). Big data and the innovation cycle. *Production and Operations Management*, *27*(9), 1642–1646. doi:10.1111/poms.12845

Li, J. P., Mirza, N., Rahat, B., & Xiong, D. (2020). Machine learning and credit ratings prediction in the age of fourth industrial revolution. *Technological Forecasting and Social Change*, *161*, 120309. doi:10.1016/j.techfore.2020.120309

Magistretti, S., Dell'Era, C., & Petruzzelli, A. M. (2019). How intelligent is Watson? Enabling digital transformation through artificial intelligence. *Business Horizons*, *62*(6), 819–829. doi:10.1016/j.bushor.2019.08.004

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, *120*, 262–273. doi:10.1016/j.jbusres.2020.07.045

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68. PMID:23074865

Moeuf, A., Lamouri, S., Pellerin, R., Tamayo-Giraldo, S., Tobon-Valencia, E., & Eburdy, R. (2020). Identification of critical success factors, risks and opportunities of Industry 4.0 in SMEs. *International Journal of Production Research*, *58*(5), 1384–1400. doi:10.1080/00207543.2019.1636323

Obschonka, M., & Audretsch, D. B. (2020). Artificial intelligence and big data in entrepreneurship: A new era has begun. *Small Business Economics*, *55*(3), 529–539. doi:10.100711187-019-00202-4

Pappas, N., Caputo, A., Pellegrini, M. M., Marzi, G., & Michopoulou, E. (2021). The complexity of decision-making processes and IoT adoption in accommodation SMEs. *Journal of Business Research*, *131*, 573–583. doi:10.1016/j.jbusres.2021.01.010

Radziwon, A., Bilberg, A., Bogers, M., & Madsen, E. S. (2014). The smart factory: Exploring adaptive and flexible manufacturing solutions. *Procedia Engineering*, *69*, 1184–1190. doi:10.1016/j.proeng.2014.03.108

Richardson, S. (2020). Cognitive automation: A new era of knowledge work? *Business Information Review*, *37*(4), 182–189. doi:10.1177/0266382120974601

Sestino, A., Prete, M. I., Piper, L., & Guido, G. (2020). Internet of Things and Big Data as enablers for business digitalization strategies. *Technovation*, *98*, 102173. doi:10.1016/j.technovation.2020.102173

Shrestha, Y. R., Krishna, V., & von Krogh, G. (2020). Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *Journal of Business Research*, *123*, 588–603. doi:10.1016/j.jbusres.2020.09.068

Siggelkow, N. (2007). Persuasion with case studies. *Academy of Management Journal*, *50*(1), 20–24. doi:10.5465/amj.2007.24160882

Sommer, L. (2015). Industrial revolution-industry 4.0: Are German manufacturing SMEs the first victims of this revolution? *Journal of Industrial Engineering and Management*, *8*(5), 1512–1532. doi:10.3926/jiem.1470

Strotmann, H. (2007). Entrepreneurial survival. *Small Business Economics*, *28*(1), 87–104. doi:10.100711187-005-8859-z

Tanev, S., & Sandstrom, G. (2019). Editorial Artificial Intelligence and Innovation Management. *Technology Innovation Management Review*, *9*(12), 3–4. doi:10.22215/timreview/1286

Trunk, A., Birkel, H., & Hartmann, E. (2020). On the current state of combining human and artificial intelligence for strategic organizational decision making. *Business Research*, 1-45.

Urbinati, A., Bogers, M., Chiesa, V., & Frattini, F. (2019). Creating and capturing value from Big Data: A multiple-case study analysis of provider companies. *Technovation*, *84*, 21–36. doi:10.1016/j.technovation.2018.07.004

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wang, L., & Alexander, C. A. (2016). Machine learning in big data. International *Journal of Mathematical. Engineering and Management Sciences*, *1*(2), 52–61.

Wang, L., Luo, G. L., Sari, A., & Shao, X. F. (2020). What nurtures fourth industrial revolution? An investigation of economic and social determinants of technological innovation in advanced economies. *Technological Forecasting and Social Change*, *161*, 120305. doi:10.1016/j.techfore.2020.120305

Wang, R., Luo, J., & Huang, S. S. (2020). Developing an artificial intelligence framework for online destination image photos identification. *Journal of Destination Marketing & Management*, *18*, 100512. doi:10.1016/j.jdmm.2020.100512

Yams, N. B., Richardson, V., Shubina, G. E., Albrecht, S., & Gillblad, D. (2020). Integrated AI and Innovation Management: The Beginning of a Beautiful Friendship. *Technology Innovation Management Review, 10*(11).

KEY TERMS AND DEFINITIONS

Artificial Intelligence: The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

Business Model: A plan for the successful operation of a business, identifying sources of revenue, the intended customer base, products, and details of financing.

Competitive Advantage: A condition or circumstance that puts a company in a favourable or superior business position.

Digital Transformation: Is the integration of digital technology into all areas of a business, fundamentally changing how you operate and deliver value to customers.

SMEs: Are businesses whose personnel numbers fall below certain limits defined in the EU recommendation 2003/361.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 6
Big Data Is Decision Science:
The Case of COVID-19 Vaccination

Jacques R. J. Bughin
Université libre de Bruxelles, Belgium

Michele Cincera
Université libre de Bruxelles, Belgium

Dorota Reykowska
NEUROHM, Poland

Rafał Ohme
WSB University in Torun, Poland

ABSTRACT

Data science has been proven to be an important asset to support better decision making in a variety of settings, whether it is for a scientist to better predict climate change for a company to better predict sales or for a government to anticipate voting preferences. In this research, the authors leverage random forest (RF) as one of the most effective machine learning techniques using big data to predict vaccine intent in five European countries. The findings support the idea that outside of vaccine features, building adequate perception of the risk of contamination, and securing institutional and peer trust are key nudges to convert skeptics to get vaccinated against COVID-19. What machine learning techniques further add beyond traditional regression techniques is some extra granularity in factors affecting vaccine preferences (twice more factors than logistic regression). Other factors that emerge as predictors of vaccine intent are compliance appetite with non-pharmaceutical protective measures as well as perception of the crisis duration.

INTRODUCTION

Big data is the handling of vast amounts of data through a flexible, mostly cloud-based IT architecture. Big data is here to stay for many reasons. The first is that digitization is data extension, with new digital

DOI: 10.4018/978-1-7998-6985-6.ch006

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

content creation growing at a rate between 40 and 60% a year. This rate means that Google, which was indexing a million pages for a few million searches in 1998, now indexing more than a trillion pages ten years later, for1.2 trillion searches a year. Second, big data can fuel new powerful machine learning techniques to uncover otherwise hidden relationships between data, that can support new powerful insights. As an example, Netflix shifted its content recommendation engine, based on customer rating, to a machine-learning algorithm fed by a large set of big data ten years ago. The tool offered a revolution to support personalized recommendations, in such a way, that now, four out of five Netflix movies and TV series viewed by Netflix subscribers originate from machine-based suggestions (Amatriain, 2013).

The seminal work by Brynjolfsson, et al. (2011) reveal that companies leveraging big data for more fine-tuned decision making, could increase their revenue productivity by more than 5%. A few years later and using a more global sample, Bughin (2016) had reached essentially the same conclusion- that is big data is usually responsible for an uplift of more than 5% in labor productivity for firms globally. Besides being used for business, big data is also important when it comes to helping major social issues, in particular when it comes to complex issues such as predictions of climate change (Bauer et al., 2021), vote intent (Mavragani and Tsagarakis, 2019), or traffic congestion (Teseng and colleagues 2018).

One another current case in point is the vaccination issue linked to the covid-19 pandemic. While the covid-19 pandemic continues unabated across the world, affecting hundreds of million people worldwide, the only serious way to eradicate the pandemic is mass- vaccination. However, large vaccination acceptance is not warranted. Social movements against vaccination have increased during the covid-19 pandemic, fueled by a series of conspiracy theories (Nguyen and Catalan, 2020). There are also possibly multiple, and interacting elements that affect vaccine intent, which make predictions of vaccine used complex—A recent poll performed by the high profile organization Pew Research in the US also emphasizes that the shape of vaccine intent is driven by « rather complex and interrelated factors » that might require more deeper data analytics to sort out a clear ranking of actions to support mass-vaccination[1].

We thus leverage machine learning to profile vaccination intent for covid-19, concentrating outside of known product factors (such as the effectiveness and safeness of the vaccine) to determine additional factors that could be boosted or alleviated to enhance fast vaccination against the Covid-19. Specifically, after aligning with recent theories of vaccine determinants to build an extensive database of behaviors and vaccine intent across 5 Continental European countries—Germany, France, Spain, Italy and Sweden, we use Random Forest (RF) techniques to provide a granular view of vaccine intent. RF is one of many machine learning techniques, but is found to generate rather accurate predictions regarding protective uptake (Bughin and Cincera, 2021)

Reaching for advanced analysis techniques has been complemented by a dedicated approach to data collection. Predicating behavior in a sensitive context such as vaccination intent in the face of the covid-19 pandemic required the usage of tools that would help decrease the natural tendency to distort declarations.

Understanding the real motives of behavior has always been a challenge for researchers. Numerous studies have shown that there is a weak correlation between declarations and behavior (Krauss, 1995) making it hard to predict behavior based solely on explicit answers. Especially when testing attitudes around sensitive topics, full of emotional load. Political correctness, post rationalizations or auto presentation needs are all important aspects that frequently influence and distort explicit, declarative answers. Additionally in the last 50 years researchers have consistently shown that most of our cognitive processes take place outside of conscious awareness and control and still they influence our perceptions, judgments and actions (Zajonc, 1968, 1980; Uleman & Bargh 1989; Bornstein & Pittman 1992; Greenwald 1992; Murphy & Zajonc 1993; Bargh 1997; Ohme, 2001; Nosek, Hawkins, & Frazier, 2011). For this reason

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in the project dedicated to understanding covid-19 vaccination intent we decided to use Response Time measurement and follow the framework of Fazio's attitude accessibility theory. Based on Fazio there are 2 components of attitudes - explicit opinion and implicit accessibility of attitude (Fazio, 2001). Attitudes that are stronger, are more accessible and therefore expressed with a shorter response time (Fazio & Williams, 1986; Fazio et al., 1989). Fazio, Powell, and Williams (1989) posited that RT might be an indicator for an attitude strength - the more quickly an attitude is expressed, the greater its strength. Stronger and more accessible attitudes have a stronger influence on behavior.

The findings are threefold. The first is the confirmation that machine learning techniques confirm a set of factors affecting the mindset of citizens with regards to vaccine intent. Those factors are in line with theories and early findings.

Second, by choosing 5 countries which were in different stages of impact, and of policy reactions to the covid-19 pandemic, we are able to sort out common and country-specific factors, according to pandemic stage and its perception by citizens. For instance, by April 2020, Sweden and Germany had suffered much lower contagion and had adopted more relaxed social distancing policies than Spain, Italy and France; hence, worries about health and home isolation that typically sort out people regarding their vaccination intent, is not emerging as such discriminant power in those countries. We also notice some differences by countries, such as media perception (lifting vaccine intent, if media exaggerate the situation).

Finally, what big data techniques bring on top of other techniques (such as multilogit regressions), is the granularity of insights; the techniques are able to point out to a set of factors that are more country- and vaccine intent type- specific, which otherwise would have remained "hidden" in the data. As an example, logit regression would show that those citizens satisfied by how their government handles the crisis have a higher propensity to get vaccinated; what big data techniques add, is that high satisfaction brings higher vaccination intent in all countries, except Italy, and that dissatisfaction correlates more with hesitancy in strict social distancing countries, than in Sweden and Germany, where dissatisfaction is only correlated with refusal to be vaccinated. This level of granularity is thus important to target the right segments of citizens.

The research reads as follows- first, we discuss the research background; then we present results of the Random Forest techniques. The last section concludes.

RESEARCH BACKGROUND

This research derives from a project at assessing citizen's attitudes, and behaviors linked with the SARS-CoV-2 pandemic, and how those aspects influence protective behavior. Bughin et al., (2020) concentrate on non-pharmaceutical protection interventions (NPIs). This research focuses on *vaccination.* The data from this survey were collected in April 2020, during the first wave of the covid-19, in France, Germany, Italy, Spain and Sweden. Data were assembled online and rely on a country representative sample for age and gender. Each country sample obeys the restriction of > 1,000 answers per country. For all questions, respondents had the choice to answer "yes, hesitant, no". Note as well that intent rates have been adjusted by response time (RT), as collected through the iCode Smart test (Ohme et al., 2020). The procedure amounts to re-center the probability of acceptance and refusal towards more hesitancy, as too quick or too long response time may indeed reduce the credibility of answers given by the respondent. About 7 points out of the 22% percent arise out of the adjusted procedure, but this should not be considered as

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

an artifact, rather an appropriate adjustment to ensure that intent is a good predictor of actual behavior (Ohme et al., 2020).

iCode™ smart test is a web-based and device agnostic technology created by NEUROHM. Apart from declarations, it captures true attitudes which are free from conformity, social and cognitive biases or wishful thinking. By measuring response time (RT) iCode™ estimates how much people hesitate when they express their opinion in a survey. Higher confidence, expressed by shorter response time indicates a well-established, internalized attitude that is more likely to drive behavior (Fazio, 2001).

For respondents the task is very simple. They are asked to evaluate if they agree with the statements presented on the screen. The answers are given on a 3 point scale (yes, hard to tell, no) and response time (RT) is measured for each answer, making it possible to collect at the same time both explicit (declarative) answers as well as implicit (indirect) attitudes.

To ensure high quality of data and eliminate test biases a calibration phase and control screen are added. Calibration precedes the test phase and consists of 3 steps:

- Familiarization with the scale
- Familiarization with the purpose of the task
- Increasing the focus on the task

Additionally a control screen is introduced to eliminate the effect of the position of the mouse on the screen. It is presented before each statement, forcing a standardized position of the mouse (the distance to the yes and no answers is always the same).

Vaccination Intent

Table 1 displays vaccination intent by countries. The share of vaccination is 65% across all countries, but with a large spread, e.g. acceptance is only 56% among Swedish citizens, and up to 74% in Spain. Furthermore, for each country, we notice that the portion of "hesitants" is larger than the share of "refusers" and more than twice in Spain for instance. The refusers segment is the largest in Germany. Vaccination intent is in line with other research, that indeed shows on average low acceptance rate in European countries and the US (see Sallam, M. (2021), Ruiz & Bell (2021) and Attwell et al. (2021). What is even more important is that intent may be too low to ensure herd immunity with certainty. E.g. in the context of a full homogenous transmission, and a reproduction rate for covid-19, at a mode of R0 at 3, herd immunity is achieved at 1-1/3= 66%. Eissler (2021), furthermore shows that for R0 above 2.5, the probability of herd immunity to be achieved, even if infection transmission is stochastic is less than 10%, thus highlighting the risk that vaccination intent today does not warrant herd immunity for sure, and thus prevent the removal of NPI as a way to limit the contagion. In such a context, we see thus as well the importance to clearly spot the drivers that would convert the hesitants to vaccination.

Conceptual Model of Vaccination Intent

We piggy back on multiple research and theories to inform about the key variables to possibly play a role on vaccine intent.

The conceptual model in mind is presented in Figure 1. The conceptual model puts asides the importance of key product attributes for acceptance. As our survey was designed before the various vaccines

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Number of Covid-19 European respondents, April 2020

Countries	Willingness to be vaccinated		
	Yes	Hesitant	Refusers
France	0.56	0.28	0.16
Germany	0.63	0.20	0.17
Italy	0.71	0.19	0.10
Spain	0.74	0.18	0.08
Sweden	0.56	0.28	0.16

Note: All variables are corrected by Response Time (RTC)

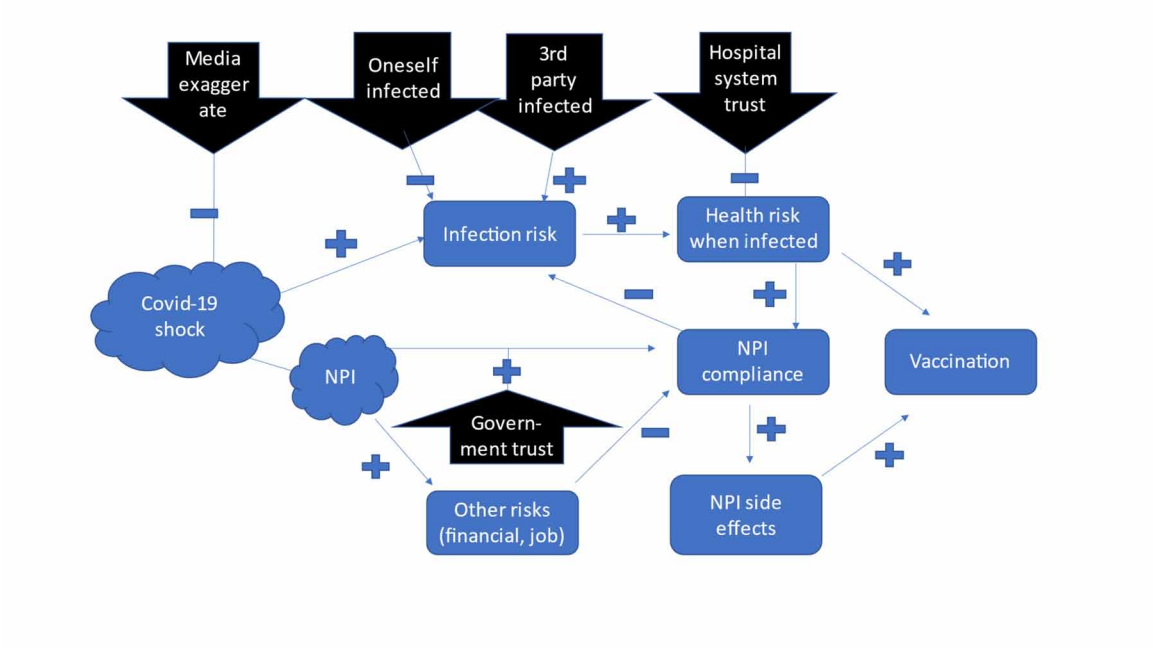
were discovered, we have based our research on the assumption that the vaccine is highly effective, with limited side effects. While not explicitly mentioned, the conceptual model is also mediated by disposition of individuals, which relates to certain types of socio-demographics, e.g. research informs of possible difference in vaccination behavior among gender, even if largely contextual. Older adults and individuals with higher education and income also show higher vaccination intention than the average, -and for a large set of vaccine uptake (Larson et al., 2014). In the case of covid-19, we state that age is especially important as the risk of lethal infection increases significantly after 60 years old. As this age matches the modal time when European citizens tend to retire from the workforce, age also impacts other financial and job sustainability risks in our conceptual model.

The conceptual model dictates that the speed of the covid-19 pandemic shock hitting most of Europe by March 2020 may have led to a worry by the population of a) being infected- and, in such a case b) of suffering health issues related to catching the virus, - a fortiori for those mistrusting the ability of the hospital care to manage the disease. The risk perception is however not uniform among citizens and depends on their indirect experience provided by the media or by the evidence of the pandemic and its health effects, among close social ties. Those who have been already infected and have recovered may adapt their risk perception, but also will have lower incentive to get vaccinated, as they have antibodies protection for a while. Also, the covid-19 shock was also accompanied by a series of more or less stringent NPI, from social distancing to quarantines. Those measures will be a fortiori more respected by those perceiving the risk on their health, as well by those obeying to social norms and /or trusting the government actions. Stringent NPI also leads to two types of negative externalities. The first, direct effect is the large economic shutdown affecting those who may be prevented to work; the second more long-term effect has been the bad-being impact of isolation, and/or too much proximity leading to family tension. The first externality would likely lead to lower NPI compliance; the second effect is a negative burden out of NPI compliance, leading to a shift towards more vaccination.

The prediction towards both NPI behavior, and vaccine intent, as a result of a pandemic shock risk perception, is based on the health belief model pioneered by Becker (1974) and discussed further in Brewer et al. (2007). Here, we also split risk in terms of getting infected, and the morbidity risk associated with being infected. We explicitly consider importance of governance and communications linked to major moral hazards, as this is often seen as a major booster of risk perception and actions, see Wachinger et al. (2013). Here we consider perceptions of citizens shaped by media communications, but also by trust in policy actions (especially NPIs) and trust in health systems quality. We finally consider the importance of social norms—e.g. health risk perception by individuals can be shaped by

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Conceptual model of vaccination



close family, friends. People also often try to influence others' to follow their behavior in the context of major disruptive events, like floating, earthquakes or pandemic (see Yang, et al. 2018). Finally, protective interventions can be pharmaceutical or not. As NPIs are known to have large negative externalities, some people would want to trade them for faster vaccination. Such tradeoff has been shown to prevail for Germany in a set of specific conjoint analyses (Bughin et al. 2021).

Table 2 provides the list of variables collected that match the model design. We have close to 50 behavioral, emotions and attitudinal questions, on top of about 10 questions linked to socio-demographics, for more than 5000 respondents. This leads to a database of more than 300,000 data points, or a « big data » set to play with. As also seen in Table 2, the considered NPIs measures include "item disinfection" (n4 and n5), or "hands cleaning" (n6), home stay and physical distancing (n2 and n3). The last two are typically the most important ones when it comes to assessing the link with health risk perception (Cho, 2020).

Regarding entities trust, we used multiple variables, e.g. for media, we look at perception of media communication exaggeration, as well as of its provision of reliable information; likewise we consider perception of government actions and information transparency. Referring to healthcare, we focus not only on how the health system manages the crisis, but also whether people perceive the health system as providing appropriate help. For instance, the perception can be that the health system has enough bed capacity for complex cases, but that the help is ineffective in intensive care, as the death rate significantly increase for people being hospitalized and getting to intensive care units.

Health worry is the largest worry expressed among citizens (62% of the total sample), but other worries as a result of shutdown were large too, e.g. job preservation was perceives as a challenge for 49% of the respondents. Isolation worries are also important, especially in countries with strict social distancing rules, during the first wave of the pandemic, like Spain and Italy. The general sense was also that

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

people were complying with NPI, even if only 55% of citizens believed that media has been providing reliable information about the pandemic. 62% believe that the health system is adequately handling this crisis, while only 55% were satisfied with how the government is handling the pandemic crisis. Last but not least, more than 7 out of 10 people responded that they were complying with physical distancing, and home quarantine. This level is high, but far from complete and is consistent with other works, e.g. Zickfeld et al. (2020).

In general, we found that differences between countries are both cultural and contextual. In a country with more social accountability like Sweden, "helping people who are vulnerable to covid-19" is an important predictor of vaccine preferences, while it is not in other countries, except Spain. In Germany, a country known, for respecting rules, compliance to NPI is more closely linked to vaccine uptake than other countries. Family tensions emerging from NPIs are naturally higher in Italy and Spain then in other countries as the latter had put less stringent distancing rules than the former.

TO BE OR NOT TO BE VACCINATED- MACHINE LEARNING INSIGHTS

This section formally tests the conceptual model.

Traditional Models

As a first step, we have run a simple univariate correlation between vaccine intent and all other variables. The exercise (not reproduced here for sake of space) shows that among all variables of Table 2, only a restricted amount of variables (about 6) exhibit significant correlation with vaccination propensity. In effect, vaccination appears to be *negatively* linked with the statement that "media exaggerate the situation with COVID-19", while it positively correlates with various indicators of health worries, (worry about one's, as well as about own children's' and older family members' health). Age, especially, 64+ older people, correlates with higher intent to get vaccinated. While the sign of the correlation fits with the literature, the univariate correlation analysis does not suggest a) any role for NPI and b) for government, healthcare and peer actions to shape vaccine preferences. Neither contextual factor such as income, nor other risk externalities beyond health, such as isolation, financial risk etc. appear to influence vaccine intent.

As a second step and further to univariate correlation analysis, we also have performed a multilogit regression on vaccine intent preferences. The high-level results are synthesized in Tables 3 and 4 for all countries pooled together. Table 3 lays out the variables that are statistically significant at 5%, while, for ease of lecture, Table 4 aggregates various indicators to fit with the conceptual model, and computes marginal impact on vaccination*.

The logistic regression is able to capture much more effect than simple one-way correlation. In fact, 17 factors out of our list come out to be statistically significant. Further, the statistically significant variables affecting vaccine intent tend to match the conceptual model, and the derived marginal effects on vaccine preferences (ref. Table 4) are about in line with other literature estimates:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Variables measured in all 5 countries

Variable	BEHAVIOR
RTC.n1	I actively encourage others to follow the restrictions and guidelines
RTC.n2	I comply with the recommendations for physical distancing
RTC.n3	I comply with the restrictions to stay home
RTC.n4	I disinfect groceries before putting them away
RTC.n5	I disinfect mail and deliveries before opening them
RTC.n6	I wash hands for 20 seconds when necessary
RTC.n7	I would like to help people who are more vulnerable to COVID-19
RTC.n8	Since COVID-19 I eat healthier
RTC.n9	Since COVID-19 I eat unhealthier
RTC.n10	Since COVID-19 I exercise less
RTC.n11	Since COVID-19 I exercise at home more
	EMOTIONS
RTC.n13	I'm worried about my financial situation
RTC.n14	I'm worried about my job situation
RTC.n15	I'm worried that our country will run out of money
RTC.n16	I'm worried that there will not be enough basic necessities in the stores
RTC.n17	I am worried about my own health
RTC.n18	I am worried about the health of my children
RTC.n19	I am worried about the health of my older family members
RTC.n20	I am worried about the health of people in my country
RTC.n21	I worry that there will be an increase in break-ins and thefts
RTC.n22	I'm worried about my children's education
RTC.n23	I am anxious about not being able to meet with friends
RTC.n24	I am worried about not being able to meet with my family
RTC.n25	I worry how living in isolation will affect me
RTC.n26	Living in isolation negatively impacts my wellbeing
	OPINIONS
RTC.n27	The COVID-19 outbreak will make society more unequal
RTC.n28	Being together all the time increases family tensions
RTC.n29	COVID-19 increases domestic violence
RTC.n30	COVID-19 will increase divorce rates
RTC.n31	COVID-19 will bring countries closer
RTC.n32	I am grateful to our essential workers
RTC.n33	I am grateful to our healthcare professionals
RTC.n34	My chance of getting COVID-19 is high
RTC.n35	Slowing the spread of COVID-19 is more important than the economy
RTC.n36	Coronavirus is dangerous for my health
RTC.n37	Media exaggerate the situation with COVID-19
RTC.n38	Media provide reliable information about the pandemic
RTC.n39	[The President] is doing a good job dealing with COVID-19
RTC.n40	I am satisfied with how my government is handling this crisis
RTC.n41	The government is doing a good job dealing with COVID-19
RTC.n42	I am satisfied with how our healthcare system is handling this crisis
RTC.n43	In the case of coronavirus infection, I will get appropriate medical help
RTC.n44	The government discloses real numbers of coronavirus infections and deaths
RTC.n45	COVID-19 reveals the best in people
RTC.n46	COVID-19 reveals the worse in people
RTC.n47	I believe we will beat COVID-19 soon
RTC.n48	People will stop following the restrictions soon

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Logit estimates, vaccination intent, April 2020, pooled for all 5 countries

VARIABLES	YES	HST	NO
26-35		0.433**	
36-49	-0.583***	0.491***	
50-64	-0.662***	0.490***	0.429*
>64	-0.519***	0.502***	
>2000 euros/month		-0.247**	
Exposed/Not to Covid		-0.345**	
Not exposed			0.420**
I would like to help people who are more vulnerable to COVID-19	0.734***	-0.660***	
Since COVID-19 I eat unhealthier			-0.448*
I'm worried about my job situation		-0.387*	
I am worried about my own health	0.316**		-0.728***
Being together all the time increases family tensions	0.409**	-0.392*	-
COVID-19 increases domestic violence		-0.276	0.455*
COVID-19 will increase divorce rates			-0.644**
I am grateful to our essential workers	0.652***		-0.794***
I am grateful to our healthcare professionals			-0.726**
My chance of getting COVID-19 is high	1.109***	-0.748***	-0.681*
Coronavirus is dangerous for my health	0.839***	-0.290*	-0.809***
Media exaggerate the situation with COVID-19	-0.830***		1.122***
Media provide reliable information about the pandemic	0.413**		-0.781***
In the case of coronavirus infection, I will get appropriate medical help	1.187***	-0.707***	-0.976***
The governments disclose the real numbers about Covid	0.494**		-1.091***

Notes: Prob>F= 0,000, pseudo R^2= 0.23

1. As said, informational factors should play a role, as at the start of the covid-19 pandemic, not so many people were contaminated, and the virus was new with large uncertainty as to its likely effects, so that people would need to base their risk perception on external clues. Two effects seem to play, but in opposite directions. Not witnessing close ties being infected plays a negative role on propensity to get vaccinated (13% lower intent of getting vaccinated), while the fact that media can provide reliable information increases the vaccine intent by the same margin.
2. Health matters have the largest direct positive effect on vaccine intent- by combining both risks of being contaminated and health issue if infected, the probability of getting vaccinated increases by 22%.
3. NPI measures do not appear to boosts vaccine intent, in contrary to NPI indirect side effects. When the effects are adverse, people seek vaccination in order to get out of their situation; but the effects are asymmetric, e.g. domestic tensions affects will to vaccinate, low risk of divorce reduces will to vaccinate.
4. (Mis-) trust in governments and healthcare quality has a significant (negative) positive impact on vaccination.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. Finally, the estimated magnitude of the effects found is in line with the literature; for instance, in the US, Kreps, et al. (2020) estimated that being exposed to a virus from known others, increases the will to get vaccinated by about 5%. In another study by Karlsson and colleagues (2021), it is estimated that increased perception of the general heath severity of the covid-19 boosts vaccine intent between 12% to 25%.

Table 4. High level effects on covid-19 vaccine intent

Effects	Marginal effects on vaccine intent
Media perception (exaggerate the situation)	-13.10%
Infection risk perception and health worry	21.80%
Medical help appropriateness	10.10%
Government trust	12.30%
NPI compliance negative side effects	10.30%
Not exposed to covid	-13.20%

Machine Learning Models

Even if already informative, traditional regression techniques may have important caveats. As an example, the regression model may have endogeneity issue; if media information can affect vaccine intent, vaccine intent can also affect how people absorb media information, as a reverse causality effect. Likewise, omitted variables might be a challenge as e.g., non-measured effects such as vaccine effectiveness, can both influence vaccine and NPI intent. Finally, a lot of the effects can be nonlinear, e.g. trust in authorities may influence risk perceptions, but more or less depending on gender, age or income (see Wachinger et al., 2013).

We thus resort to deep learning approaches, that allow for providing more robust predictions given the restrictions of the above, but also because we want to analyze non-linear effects, that could be of high importance to reveal nudges to reduce the portion of refusers and hesitant. In particular, we have tested multiple learning techniques, from GVM to neural networks, etc. As we have found that the techniques provide same qualitative results, we herewith report on the Random Forest (RF) technique, as the technique is the most often used as big data technique related to pandemics; for example, Random Forest techniques have exhibited superior predictive power for H5N1 influenza outbreaks see Kane et al. (2014), and recently for covid-19 infections (Yadav et al., 2020; Yeşilkanat, 2020). Here, we have configured the Random Forest algorithm with 5,000 trees in the forest, including statements and demographics, with all variables being tried at each split, in order to understand hidden complex tree structure to predict vaccination intent. The model was run for each country specifically and we present the results by country in order to gauge country commonalties and specificities. The analysis also looks at the three key segments of vaccine preferences (acceptors (YES), refusers (NO) and hesitants (HST)). Table 5 lays out the predictive power of the RF algorithm per country based on root mean square error (RMSE). Tables 6 to 10 provide results for each of the five countries in terms of contribution to accuracy (in points of percentage), and by segments of vaccine preferences.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

If the RMSE remains large, it is nevertheless in the range of 0.25-0.5 for RF models. This level of accuracy supports the fact that the model can relatively predict the data accurately. Further, we have observed that the RMSE of RF techniques is lower than with the logistic regression. This is especially the case for hesitant and refusers, --two segments that are problematic, and must be nudged to change their mind if one wants to ensure herd immunity.

Table 5. Predictive performance comparison (10-fold cross-validation) of Random Forest models

Model	YES	HST	NO
	RMSE		
Germany	0.4205	0.3954	0.3299
France	0.4581	0.4374	0.3492
Italy	0.4223	0.3804	0.2886
Spain	0.4158	0.3720	0.2692
Sweden	0.4636	0.4380	0.3511
Model	**YES**	**HST**	**NO**
	RMSE RF> RMSE logit		
Germany	YES		YES
France		YES	YES
Italy		YES	YES
Spain		YES	
Sweden	YES	YES	

One confirms, from those tables, that RF generates a broader list of variables (from 31 for Spain to up to 41 in France) that affect vaccine preferences, than the list of 17 variables that has emerged out of traditional techniques. We also note that the difference seems to be that RF provides a longer-tail of variables that are often either country, or segment specific.

Common factors

Regarding common factors, two elements converge with the logistic model predictions. Health risk perceptions are largest contributors to vaccine preferences prediction accuracy. “Being worried about own health”, and “coronavirus is dangerous for my health”, clearly discriminate between acceptor of vaccine or not, in all countries. Indirect perception linked to media further shapes the type of non-acceptance, as media exaggeration is a consistent tag line for refusers (versus hesitant), while media reliability of information emerges consistently across countries as a universal driver of acceptance (versus hesitation).

We also witness the critical importance of institutional trust, a factor that has been noted as having one of the most substantial impacts on risk perception of natural hazards; see Yadav et al., 2020, for the covid-19 as well as Terpstra (2011) for flood impact. Trust (as well as satisfaction) in healthcare/government (actions) both play a role, with the trust in healthcare generally slightly more predictive for vaccination, than trust in government actions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

On top of the logistic regression, RF provides additional elements that were hidden by the technique of regression, both because they affect segment preferences asymmetrically, and/or because the effects of those factors were absorbed and tied to other variables:

1. The first is peer trust, acting on top of institutional trust. This *peer trust* is measured by two markers, « covid reveals the best/worst of people », and « The Covid crisis will bring countries closer ». The lack of peer trust makes vaccination less likely, as it was also shown to reduce NPI compliance during the covid-19, see Mehari (2020). As peer trust is correlated with government trust, the effect got blurred into government trust in the regression analysis;
2. The second element is the *balance between economics and health*, as preference for the later versus the former leads to a clear separation between refusers and acceptors of vaccine uptake. This effect has remained hidden in the regression technique, as overweight to health correlates strongly with health worries, while the effect on vaccination depends as well on other elements such as age.
3. The third is that, as per the conceptual model, *compliance with NPI measures predicts vaccine intent*, while indirect side effects of those measures, especially bad-being linked to isolation, also leads to more vaccine intent. Note that the isolation effect is large enough to have refusers switch to become more hesitant, but still not yet fully convinced to be vaccinated.
4. The fourth element is *the expectation of the duration of the crisis*. An expectation that the crisis will not last long reduces acceptors in favor of hesitant for all countries under analysis.

Specific Factors

From the list, a few specific factors finally emerge that are either country or segment dependent.

In general country differences can be expected, and in particular because of the different measures undertaken by the authorities. (Full country comparison can be found in Tables 11, 12 and 13 in appendix) For example, while third parties' health worries are positively associated with higher propensity to get the vaccine, third party's most concern is about children in Germany, while it is more about older family members in Italy. Italy in fact, was put in a stricter lockdown than Germany, due to pandemic not being in control, and large mortality happening among older population. This lockdown made the Italians more at risk to eat unhealthy and not exercise; those able to stick to healthy practices were thus less inclined to feel bad about lockdown, with higher reluctance to vaccinate. Isolation and inability to meet friends were not seen as tangible risks in Sweden given that no strict lockdown was prevailing in the first wave of the pandemic. Lockdown in form of quarantine when sick was the only hazard Swedish citizens had to face, as social distancing was not mandatory for citizens; thus being confronted with being contaminated was a way to reduce anti-vaccination in the country. Segment differences are more visible with RF than with traditional regressions, in general, because the effects play non-linearly. Age does not play a role for hesitant, but more between acceptors and those reluctant to get vaccinated, and especially for all age brackets in France. Low income plays a negative role on intent, especially for Italy, as an example.

In general and as the various Tables 6-10 show, the extra factors emerging from the RF exercise (about 20 more factors) are essentially effects that are clustering by segment preferences and/or countries, and are typically known correlates of other main factors (e.g. low income, low education correlate with lower citizen trust).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. "Acceptors", "Hesitant" and "Refusers": Factors in common or country specific - Germany

Factors	#	YES	HST	NO
I actively encourage others to follow the restrictions and guidelines	3	17.131	1.344	6.477
I am worried about my own health	3	4.597	1.185	2.484
Coronavirus is dangerous for my health	3	20.196	1.376	8.585
Media exaggerate the situation with COVID-19	3	5.507	1.144	14.41
Media provide reliable information about the pandemic	3	2.976	1.183	1.578
The government is doing a good job dealing with COVID-19	3	4.752	1.319	1.799
# of children 0	3	2.747	1.254	1.571
I comply with the restrictions to stay home	2	3.107	1.92	
I wash hands for 20 seconds when necessary	2	2.586	2.495	
COVID-19 increases domestic violence	2	2.604		1.212
I am anxious about not being able to meet with friends	2	3.713		1.543
I am grateful to our healthcare professionals	2	4.457	3.449	
I am satisfied with how my government is handling this crisis	2	5.599	4.834	
I am satisfied with how our healthcare system is handling this crisis	2	2.728	2.157	
In case of a coronavirus infection I will get appropriate medical help	2	3.858		1.24
The government discloses real numbers of coronavirus infections and deaths	2	4.062		1.171
The restrictions caused by COVID-19 will continue for about a month	2	4.085		1.691
politics Other	2	3.148	1.507	
age 26-35	1		1.5	
I would like to help people who are more vulnerable to COVID-19	1			1.226
I am worried about my financial situation	1			1.147
I am worried that our country will run out of money	1		1.552	
I am worried that there will not be enough basic necessities in the stores	1		1.811	
I am worried about the health of my children	1		1.586	
I am worried about the health of my older family members	1		1.565	
I am worried about my children s education	1		1.75	
COVID-19 will increase divorce rates	1			1.137
I am worried about not being able to meet with my family	1			1.486
Living in isolation negatively impacts my wellbeing	1			1.322
COVID-19 will bring countries closer	1	2.683		
Slowing the spread of COVID-19 is more important than the economy	1			1.218
[PRESIDENT] is doing a good job dealing with COVID-19	1	2.63		
COVID-19 reveals the worse in people	1		2.259	
People will stop following the restrictions soon	1			1.174
The restrictions caused by COVID-19 will continue at least until the fall	1			1.131

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. "Acceptors", "Hesitant" and "Refusers": Factors in common or country specific – France

Factors	#	H2T	NO	YES
I disinfect groceries before putting them away	3	1.51	3.015	6.175
In case of a coronavirus infection I will get appropriate medical help	3	1.426	2.601	4.301
Don't associate with politics	3	1.881	1.866	5.621
I wash hands for 20 seconds when necessary	2		2.558	3.296
I would like to help people who are more vulnerable to COVID-19	2	1.517	1.89	
Since COVID-19 I exercise less	2	1.532		3.313
I am worried about my own health	2		3.089	6.441
Being together all the time increases family tensions	2	1.535	1.865	
I am anxious about not being able to meet with friends	2	1.429		3.486
Coronavirus is dangerous for my health	2		3.328	5.738
Media exaggerate the situation with COVID-19	2	1.429	3.785	
I am satisfied with how my government is handling this crisis	2	1.527		4.175
I am satisfied with how our healthcare system is handling this crisis	2	1.581		3.499
The government discloses real numbers of coronavirus infections and deaths	2	1.444		3.574
People will stop following the restrictions soon	2	1.478	1.89	
The restrictions caused by COVID-19 will continue at least until the fall	2	1.711		4.218
age 36-49	1		2.354	
age 50-64	1		1.718	
age >64	1			3.977
I disinfect mail and deliveries before opening them	1			4.528
Since COVID-19 I eat more healthy	1	1.574		
Since COVID-19 I exercise at home more	1			3.071
I am worried about my financial situation	1	1.487		
I am worried about the health of my older family members	1		1.909	
COVID-19 will increase divorce rates	1	1.597		
I worry how living in isolation will affect me	1	1.597		
Living in isolation negatively impacts my wellbeing	1	1.523		
COVID-19 will bring countries closer	1			6.177
I am grateful to our essential workers	1		1.848	
Media provide reliable information about the pandemic	1			4.336
COVID-19 reveals the best in people	1			3.578
COVID-19 reveals the worse in people	1		1.904	
I believe we will beat COVID-19 soon	1		2.028	
The restrictions caused by COVID-19 will continue for about a month	1	1.562		
Middle school	1		1.673	
Vocational	1		1.924	
income <20000€	1		1.961	
Retired	1			5.051
politics Right	1			3.577
politics Other	1		1.81	
Male	1	1.456		

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. "Acceptors", "Hesitant" and "Refusers": Factors in common or country specific – Italy

Factors	#	H2T	NO	YES
Since COVID-19 I exercise at home more	3	1.254	1.023	2.883
I am worried about my own health	3	1.051	1.168	3.03
Coronavirus is dangerous for my health	3	1.694	3.883	15.414
Media provide reliable information about the pandemic	3	1.317	1.433	5.733
In case of a coronavirus infection I will get appropriate medical help	3	1.553	1.369	5.343
income <20000€	3	1.439	1.493	2.67
age 36-49	2		1.147	3.072
I actively encourage others to follow the restrictions and guidelines	2		4.869	4.075
Since COVID-19 I exercise less	2	1.056	1.026	
I am worried about the health of my older family members	2	1.063		2.595
Slowing the spread of COVID-19 is more important than the economy	2	1.218		2.816
[PRESIDENT] is doing a good job dealing with COVID-19	2	1.37		3.08
The government is doing a good job dealing with COVID-19	2	1.246		3.339
The government discloses real numbers of coronavirus infections and deaths	2	1.109		4.001
I believe we will beat COVID-19 soon	2		1.864	2.422
income >20000€	2	1.544		2.451
income don't want to answer	2		1.348	2.566
I comply with the restrictions to stay home	1		1.039	
I disinfect groceries before putting them away	1		1.02	
I wash hands for 20 seconds when necessary	1		1.455	
I would like to help people who are more vulnerable to COVID-19	1	1.131		
Since COVID-19 I eat more healthy	1	1.106		
Since COVID-19 I eat more unhealthy	1		1.073	
I am worried that our country will run out of money	1		1.548	
Being together all the time increases family tensions	1			2.679
COVID-19 will increase divorce rates	1		1.146	
I am worried about not being able to meet with my family	1	1.049		
I worry how living in isolation will affect me	1	1.191		
Living in isolation negatively impacts my wellbeing	1	1.241		
Media exaggerate the situation with COVID-19	1		1.259	
I am satisfied with how our healthcare system is handling this crisis	1	1.125		
Vocational	1		1.534	
High school	1			2.424
# of children 2	1			2.601
politics Left	1			3.113
politics Other	1	1.094		
Don't associate with politics	1		1.029	

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9. "Acceptors", "Hesitant" and "Refusers": Factors in common or country specific – Spain

Factors	#	H2T	NO	YES
I would like to help people who are more vulnerable to COVID-19	3	1.067	0.5574	2.967
I am worried that there will not be enough basic necessities in the stores	3	1.172	0.6056	3.035
I am worried about my own health	3	1.245	0.5962	3.148
Being together all the time increases family tensions	3	1.398	0.6332	4.277
I am worried about not being able to meet with my family	3	1.061	0.679	3.564
I worry how living in isolation will affect me	3	1.071	0.6362	2.667
Slowing the spread of COVID-19 is more important than the economy	3	1.01	0.6841	3.138
Coronavirus is dangerous for my health	3	1.236	0.6955	4.37
I am satisfied with how our healthcare system is handling this crisis	3	1.19	0.5885	3.314
Don't associate with politics	3	1.316	0.6062	5.168
Since COVID-19 I eat more healthy	2	1.079		2.739
I worry that there will be an increase in break-ins and thefts	2	1.067		2.759
I am grateful to our essential workers	2		0.9236	5.093
I am grateful to our healthcare professionals	2		0.8812	3.935
Media exaggerate the situation with COVID-19	2		1.0383	3.069
COVID-19 reveals the best in people	2	1.179		2.525
COVID-19 reveals the worse in people	2	1.032	0.581	
The restrictions caused by COVID-19 will continue at least until the fall	2		0.6249	2.762
The restrictions caused by COVID-19 will continue for about a month	2	1.21		2.515
I disinfect groceries before putting them away	1		0.6347	
I disinfect mail and deliveries before opening them	1		0.7	
Since COVID-19 I exercise less	1	1.015		
The COVID-19 outbreak will make society more unequal	1	1.048		
I am worried about the health of my older family members	1		0.613	
I am anxious about not being able to meet with friends	1		0.6025	
Living in isolation negatively impacts my wellbeing	1	1.066		
COVID-19 will bring countries closer	1	1.02		
My chance of getting COVID-19 is high	1			2.676
In case of a coronavirus infection I will get appropriate medical help	1		0.5866	
People will stop following the restrictions soon	1			2.813
income >20000€	1	1.004		

We conclude by looking at the hesitants, as this segment is twice as large as the refusers, and is the one that one would hope to influence to convert to vaccination as a way to reach herd immunity. Logit regression results shown in Table 3 suggest that hesitants are a category of less assertive refusers, except that family tension (but not violence) leads them to convert to accept vaccine. What RF does show is that you can find better nudges to make them convert. First, this is linked to the reality of the pandemic, that is pandemic will not stop easily and will last long without enough vaccinated to reach immunity. Second,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. "Acceptors", "Hesitant" and "Refusers": Factors in common or country specific - Sweden

Factors	#	H2T	NO	YES
I would like to help people who are more vulnerable to COVID-19	3	1.628	0.9341	3.305
I am worried about my own health	3	1.521	1.1974	7.91
In case of a coronavirus infection I will get appropriate medical help	3	1.676	1.0214	6.545
The government discloses real numbers of coronavirus infections and deaths	3	1.498	1.0189	3.421
People will stop following the restrictions soon	3	1.595	1.0771	3.231
Since COVID-19 I exercise at home more	2		0.9342	2.947
I am worried about my financial situation	2	1.474	0.9647	
I am worried that our country will run out of money	2	1.534	1.0807	
Living in isolation negatively impacts my wellbeing	2	1.488		3.331
Slowing the spread of COVID-19 is more important than the economy	2		1.0827	3.512
Coronavirus is dangerous for my health	2		1.289	5.605
Media provide reliable information about the pandemic	2	1.506		3.954
[PRESIDENT] is doing a good job dealing with COVID-19	2	1.575		4.894
I am satisfied with how my government is handling this crisis	2		0.909	3.079
The government is doing a good job dealing with COVID-19	2	1.581		3.712
I am satisfied with how our healthcare system is handling this crisis	2	1.7		3.264
COVID-19 reveals the best in people	2	1.491		3.67
COVID-19 reveals the worse in people	2	1.547	0.909	
Quarantine no	2	1.52		3.218
age 26-35	1		1.0788	
age 36-49	1		1.1173	
age >64	1			6.116
I comply with the recommendations for physical distancing	1		0.9211	
I comply with the restrictions to stay home	1		1.0453	
The COVID-19 outbreak will make society more unequal	1		0.9766	
I am worried about the health of my older family members	1		1.0895	
COVID-19 increases domestic violence	1			3.007
COVID-19 will increase divorce rates	1	1.55		
I am worried about not being able to meet with my family	1		1.117	
COVID-19 will bring countries closer	1			4.05
My chance of getting COVID-19 is high	1			3.176
Media exaggerate the situation with COVID-19	1		1.5861	
I believe we will beat COVID-19 soon	1	1.477		
The restrictions caused by COVID-19 will continue at least until the fall	1	1.483		
The restrictions caused by COVID-19 will continue for about a month	1	1.553		
Don't associate with politics	1	1.469		

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

it is important to demonstrate better institutional trust, with much more appropriate and educated media information. This effect of trust has been already shown to have a substantial effect on risk perception in many natural hazards. We add covid-19 pandemic to this list.

CONCLUSION

This article has used big data techniques, in particular Random Forest technique, to provide a more robust baseline regarding the covd-19 vaccine preferences. As we hypothesized that when it comes to vaccine decisions, a large set of factors interact and possibly with asymmetry, we demonstrate that RF provides a much more granular view as to the factors affecting vaccine choices. Good news is that the techniques do not necessarily conflict. The same common, first-order factors emerge as predictors in all techniques, but RF clearly provides a more complete view.

In general, the research is a first to show that compliance to non-pharma protective measures is already a good marker of the citizen mindset for vaccine acceptance, as is the perception of how long will the crises last. As per other literature, the key nudge remains to objective the real risk of the pandemic, and make people also aware of the fact that a pandemic is only to be stopped by limiting self- defeating behavior, like reluctance to, or free-riding on protective measures.

ACKNOWLEDGMENT

Special acknowledgement should be given to NEUROHM for initiating and supervising the projects and SYNO for collecting the data.

REFERENCES

Amatriain, X. (2013) *Beyond Data: from user information to business value through personalized recommendations and consumer science*. Paper presented at the CIKM'13, San Francisco, CA.

Attwell, K., Lake, J., Sneddon, J., Gerrans, P., Blyth, C., & Lee, J. (2021). Converting the maybes: Crucial for a successful COVID-19 vaccination strategy. *PLoS One*, *16*(1), e0245907. doi:10.1371/journal.pone.0245907 PMID:33471821

Bargh, J. A. (1997). Automaticity of Everyday Life. *Advances in Social Cognition*, *10*, 1–61.

Bauer, P., Stevens, B., & Hazeleger, W. (2021). A digital twin of Earth for the green transition. *Nature Climate Change*, *11*(2), 80–83. doi:10.103841558-021-00986-y

Bornstein, R. F., & Pittman, T. S. (1992). *Perception Without Awareness*. The Guilford Press.

Becker, M. H. (1974). The health belief model and sick role behavior. *Health Education Monographs*, *2*(4), 409–419. doi:10.1177/109019817400200407

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D., & Weinstein, N. D. (2007). Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychology*, *26*(2), 136–145. doi:10.1037/0278-6133.26.2.136 PMID:17385964

BrynjolfssonE.HittL.KimH. (2011). Strength in numbers: how does data-driven decision making affect firm performance? https://ssrn.com/abstract=1819486 doi:10.2139srn.1819486

Bughin, J. (2016). Big data, Big bang? *Journal of Big Data*, *3*(1), 1–14. doi:10.118640537-015-0014-3

BughinJ.CinceraM. (2021). *How Institutional Actions before Vaccine Affects Vaccination Intention Later: Prediction Via Machine Learning*. https://ssrn.com/abstract=3824550 doi:10.2139srn.3824550

BughinJ.CinceraM.ReykowskaD.ŻyszkiewiczM.OhmeR. (2021). COVID-19 Endemism and the Control Skeptics. https://ssrn.com/abstract=3785230 doi:10.2139srn.3785230

Bughin, J., Cincera, M., Reykowska, D., Żyszkiewicz, M., & Ohme, R. (2020). *Perceptive risk clusters of European citizens and NPI compliance in face of the covid-19 pandemics.* Working Papers ECARES 2020-51, ULB Université libre de Bruxelles.

Cho, S. W. (2020). Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. *The Econometrics Journal*, *23*(3), 323–334. doi:10.1093/ectj/utaa025

Eissler, M. (2021). Probability that an infection like Covid-19 stops without reaching herd immunity, calculated with a stochastic agent-based model. medRxiv, 2020-12.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*(2), 115–141. doi:10.1080/02699930125908

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *The Journal of Consumer Research*, *16*(3), 280–288. doi:10.1086/209214

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238. doi:10.1037/0022-3514.50.2.229 PMID:3701576

Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude–perception and attitude–behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology*, *51*(3), 505–514. doi:10.1037/0022-3514.51.3.505 PMID:3761146

Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *The American Psychologist*, *47*(6), 766–779. doi:10.1037/0003-066X.47.6.766 PMID:1616174

Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, *15*(1), 276. doi:10.1186/1471-2105-15-276 PMID:25123979

Karlsson, L. C., Soveri, A., Lewandowsky, S., Karlsson, L., Karlsson, H., Nolvi, S., Karuki, M., Lindfelt, M., & Antfolk, J. (2021). Fearing the disease or the vaccine: The case of COVID-19. *Personality and Individual Differences*, *172*, 110590. doi:10.1016/j.paid.2020.110590 PMID:33518869

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kraus, S. J. (1995). Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature. *Personality and Social Psychology Bulletin*, *21*(1), 58–75. doi:10.1177/0146167295211007

Kreps, S., Prasad, S., Brownstein, J. S., Hswen, Y., Garibaldi, B. T., Zhang, B., & Kriner, D. L. (2020). Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA Network Open*, *3*(10), e2025594–e2025594. doi:10.1001/jamanetworkopen.2020.25594 PMID:33079199

Larson, H., Jarrett, C., Eckersberger, E., Smith, D., & Paterson, P. (2014). Understanding vaccine hesitancy around vaccines from a global perspective: A systematic review of published literature, 2007-2012. *Vaccine*, *32*(19), 2150–2159. doi:10.1016/j.vaccine.2014.01.081 PMID:24598724

Mavragani, A., & Tsagarakis, K. P. (2019). Predicting referendum results in the Big Data Era. *Journal of Big Data*, *6*(1), 1–20. doi:10.118640537-018-0166-z

Mehari, Y. (2020). *The Role of Social Trust in Citizen Mobility During COVID-19*. https://ssrn.com/abstract=3607668[REMOVED HYPERLINK FIELD] doi:10.2139srn.3607668

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*(5), 723–739. doi:10.1037/0022-3514.64.5.723 PMID:8505704

Nguyen, A., & Catalan, D. (2020). Digital mis/disinformation and public engagement with health and science controversies: Fresh perspectives from Covid-19. *Media and Communication*, *8*(2), 323–328. doi:10.17645/mac.v8i2.3352

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152–159. doi:10.1016/j.tics.2011.01.005 PMID:21376657

Ohme, R. (2001). The implicit conditioning of consumer attitudes: Logo substitution effect. Implicit Processes in Cognition, 32(1).

Ohme, R., Matukin, M., & Wicher, P. (2020). Merging Explicit Declarations With Implicit Response Time to Better Predict Behavior. In V. Chkoniya, A. O. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 427–448). IGI Global. doi:10.4018/978-1-7998-3115-0.ch023

Ruiz, J. B., & Bell, R. A. (2021). Predictors of intention to vaccinate against COVID-19: Results of a nationwide survey. *Vaccine*, *39*(7), 1080–1086. doi:10.1016/j.vaccine.2021.01.010 PMID:33461833

Sallam, M. (2021). COVID-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates. *Vaccines*, *9*(2), 160. doi:10.3390/vaccines9020160 PMID:33669441

Terpstra, T. (2011). Emotions, trust, and perceived risk: Affective and cognitive routes to floofd preparedness. *Risk Analysis*, *31*(10), 1658–1675. doi:10.1111/j.1539-6924.2011.01616.x PMID:21477090

Uleman J. S. & Bargh J. A. (1989). *Unintended thought*. New York: Guilford Press.

Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox: Implications for governance and communication of natural hazards. *Risk Analysis*, *33*(6), 1049–1065. doi:10.1111/j.1539-6924.2012.01942.x PMID:23278120

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yadav, M., Perumal, M., & Srinivas, M. (2020). Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos, Solitons, and Fractals*, *139*, 110050. doi:10.1016/j.chaos.2020.110050 PMID:32834604

Yang, C., Dillard, J., & Li, R. (2018). Understanding fear of Zika: Personal, interpersonal, and media Influences. *Risk Analysis*, *38*(12), 2535–2545. doi:10.1111/risa.12973 PMID:29392760

Yeşilkanat, C. M. (2020). Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine-learning algorithm. *Chaos, Solitons, and Fractals*, *140*, 110210. doi:10.1016/j.chaos.2020.110210 PMID:32843823

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. doi:10.1037/h0025848

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *The American Psychologist*, *35*(2), 151–175. doi:10.1037/0003-066X.35.2.151

Zickfeld, J. H., Schubert, T. W., Herting, A. K., Grahe, J., & Faasse, K. (2020). Correlates of health-protective behavior during the initial days of the COVID-19 outbreak in Norway. *Frontiers in Psychology*, *11*, 11. doi:10.3389/fpsyg.2020.564083 PMID:33123045

ADDITIONAL READING

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324

Jovančević, A., & Milićević, N. (2020). Optimism-pessimism, conspiracy theories and general trust as factors contributing to COVID-19 related behavior–A cross-cultural study. *Personality and Individual Differences*, *167*, 110216. doi:10.1016/j.paid.2020.110216 PMID:32834280

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., & El-Mohandes, A. (2020). A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*, 1–4. doi:10.103841591-020-1124-9 PMID:33082575

Neumann-Bohme, S., Varghese, N. E., Sabat, I., Barros, P. P., Brouwer, W., Van Exel, J., & Stargardt, T. (2020). Once we have it, will we use it? A European survey on willingness to be vaccinated against COVID-19. *The European Journal of Health Economics*, *21*(7), 977–982. doi:10.100710198-020-01208-6 PMID:32591957

KEY TERMS AND DEFINITION

Attitudes: Can be defined as an organized predisposition to respond in a favorable or unfavorable manner toward a specified class of objects.

Big Data: Big and versatile data sets, that are hard to analyze using standard analytical methods. Analysis of big data enables to identify new information and gain knowledge by finding relationships between variables not visible before.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

iCode™: A unique tool created by NEUROHM to test response time (RT) in the field of brands, products, commercials, motives, emotional experience, sensory testing, services, satisfaction, or personal assessment.

Machine Learning: An area of artificial intelligence dedicated to algorithms that improve automatically through experience obtained via data exposure. Machine learning algorithms build a mathematical model from sample data, called a training set, to forecast or make decisions without being explicitly programmed by humans for that purpose. Machine learning algorithms are used in many different applications, such as image recognition, speech recognition, data analysis and classification, anti-spam etc.

Random Forest: Random forest is a supervised learning algorithm that works as an ensemble learning method based on building a range of regression trees that are then averaged out to compose the final forest. From each tree, the smallest root mean square error determines the top of the tree and recursively creates a full tree. Prediction uses the average of the response variable in each leaf of the tree. The advantages of Random Forests are: they are robust to outliers, effective with nonlinear data, and have lower risk of overfitting compared to decision trees.

Response Time (RT): Time necessary to produce an answer. Method frequently used by psychologists to measure, for example, attitude strength.

ENDNOTES

* Note that only statistically significant regression coefficients are displayed and added as marginal impact to vaccine effects. We nevertheless have taken aside significant country fixed effects as well as demographic effects

1 Intent to Get a COVID-19 Vaccine Rises to 60% as Confidence in Research and Development Process Increases | Pew Research Center

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Table 11. Factors in common and specific to countries: "Acceptors"

Factors	#	DE	FR	IT	SE	SP
I am worried about my own health	5	1	1	1	1	1
Coronavirus is dangerous for my health	5	1	1	1	1	1
Media provide reliable information about the pandemic	4	1	1	1	1	
I am satisfied with how our healthcare system is handling this crisis	4	1	1		1	1
In case of a coronavirus infection I will get appropriate medical help	4	1	1	1	1	
The government discloses real numbers of coronavirus infections and deaths	4	1	1	1	1	
Since COVID-19 I exercise at home more	3		1	1	1	
COVID-19 will bring countries closer	3	1	1		1	
Slowing the spread of COVID-19 is more important than the economy	3			1	1	1
[PRESIDENT] is doing a good job dealing with COVID-19	3	1		1	1	
I am satisfied with how my government is handling this crisis	3	1	1		1	
The government is doing a good job dealing with COVID-19	3	1		1	1	
COVID-19 reveals the best in people	3		1		1	1
age >64	2		1		1	
I actively encourage others to follow the restrictions and guidelines	2	1		1		
I wash hands for 20 seconds when necessary	2	1	1			
I would like to help people who are more vulnerable to COVID-19	2				1	1
Being together all the time increases family tensions	2			1		1
COVID-19 increases domestic violence	2	1			1	
I am anxious about not being able to meet with friends	2	1	1			
I am grateful to our healthcare professionals	2	1				1
My chance of getting COVID-19 is high	2				1	1
Media exaggerate the situation with COVID-19	2	1				1
People will stop following the restrictions soon	2				1	1
The restrictions caused by COVID-19 will continue at least until the fall	2		1			1
The restrictions caused by COVID-19 will continue for about a month	2	1				1
Don't associate with politics	2		1			1
I comply with the restrictions to stay home	1	1				
I disinfect groceries before putting them away	1		1			
I disinfect mail and deliveries before opening them	1		1			
Since COVID-19 I eat more healthy	1					1
Since COVID-19 I exercise less	1		1			
I am worried that there will not be enough basic necessities in the stores	1					1
I am worried about the health of my older family members	1			1		
I worry that there will be an increase in break-ins and thefts	1					1
I am worried about not being able to meet with my family	1					1
I worry how living in isolation will affect me	1					1
Living in isolation negatively impacts my wellbeing	1				1	
I am grateful to our essential workers	1					1
I believe we will beat COVID-19 soon	1			1		
High school	1			1		
income <20000€	1			1		
income >20000€	1			1		
income don't want to answer	1			1		
# of children 0	1	1				
# of children 2	1			1		
Retired	1		1			
politics Left	1			1		
politics Right	1		1			
politics Other	1	1				
Quarantine no	1				1	

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 12. Factors in common and specific to countries: "Hesitant"

Factors	#	DE	FR	IT	SE	SP
I would like to help people who are more vulnerable to COVID-19	5	1	1	1	1	1
Living in isolation negatively impacts my wellbeing	5	1	1	1	1	1
I am worried about my own health	4	1		1	1	1
I am satisfied with how our healthcare system is handling this crisis	4		1	1	1	1
In case of a coronavirus infection I will get appropriate medical help	4	1	1	1	1	
The government discloses real numbers of coronavirus infections and deaths	4	1	1	1	1	
The restrictions caused by COVID-19 will continue for about a month	4	1	1		1	1
Since COVID-19 I eat more healthy	3		1	1		1
Since COVID-19 I exercise less	3		1	1		1
I am worried about my financial situation	3	1	1		1	
COVID-19 will increase divorce rates	3	1	1		1	
I am worried about not being able to meet with my family	3	1		1		1
I worry how living in isolation will affect me	3		1	1		1
Slowing the spread of COVID-19 is more important than the economy	3	1		1		1
Coronavirus is dangerous for my health	3	1		1		1
Media provide reliable information about the pandemic	3	1		1	1	
The government is doing a good job dealing with COVID-19	3	1		1	1	
People will stop following the restrictions soon	3	1	1		1	
The restrictions caused by COVID-19 will continue at least until the fall	3	1	1		1	
Don't associate with politics	3		1		1	1
Being together all the time increases family tensions	2		1			1
I am anxious about not being able to meet with friends	2	1	1			
Media exaggerate the situation with COVID-19	2	1	1			
COVID-19 reveals the best in people	2				1	1
COVID-19 reveals the worse in people	2				1	1
income >20000€	2			1		1
I actively encourage others to follow the restrictions and guidelines	1	1				
I disinfect groceries before putting them away	1		1			
Since COVID-19 I exercise at home more	1			1		
I am worried that our country will run out of money	1				1	
I am worried that there will not be enough basic necessities in the stores	1					1
The COVID-19 outbreak will make society more unequal	1					1
I am worried about the health of my older family members	1			1		
I worry that there will be an increase in break-ins and thefts	1					1
COVID-19 increases domestic violence	1	1				
COVID-19 will bring countries closer	1					1
[PRESIDENT] is doing a good job dealing with COVID-19	1			1		
[PRESIDENT] is doing a good job dealing with COVID-19	1				1	
I am satisfied with how my government is handling this crisis	1		1			
I believe we will beat COVID-19 soon	1				1	
income <20000€	1			1		
# of children 0	1	1				
politics Other	1			1		
Quarantine no	1				1	
Male	1		1			

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 13. Factors in common and specific to country: "Refusers"

Factors	#	DE	FR	IT	SE	SP
I am worried about my own health	5	1	1	1	1	1
Coronavirus is dangerous for my health	5	1	1	1	1	1
Media exaggerate the situation with COVID-19	5	1	1	1	1	1
I am worried about the health of my older family members	4	1	1		1	1
In case of a coronavirus infection I will get appropriate medical help	4		1	1	1	1
COVID-19 reveals the worse in people	4	1	1		1	1
age 36-49	3		1	1	1	
I comply with the restrictions to stay home	3	1		1	1	
I disinfect groceries before putting them away	3		1	1		1
I wash hands for 20 seconds when necessary	3	1	1	1		
I would like to help people who are more vulnerable to COVID-19	3		1		1	1
I am worried that our country will run out of money	3	1		1	1	
Don't associate with politics	3		1	1		1
age 26-35	2	1			1	
I actively encourage others to follow the restrictions and guidelines	2	1		1		
Since COVID-19 I exercise at home more	2			1	1	
I am worried that there will not be enough basic necessities in the stores	2	1				1
Being together all the time increases family tensions	2		1			1
I am worried about not being able to meet with my family	2				1	1
I am grateful to our essential workers	2		1			1
I am grateful to our healthcare professionals	2	1				1
Slowing the spread of COVID-19 is more important than the economy	2				1	1
Media provide reliable information about the pandemic	2	1		1		
I am satisfied with how my government is handling this crisis	2	1			1	
I am satisfied with how our healthcare system is handling this crisis	2	1				1
I believe we will beat COVID-19 soon	2		1	1		
People will stop following the restrictions soon	2		1		1	
Vocational	2		1	1		
income <20000€	2		1	1		
politics Other	2	1	1			
age 50-64	1		1			
I comply with the recommendations for physical distancing	1				1	
I disinfect mail and deliveries before opening them	1					1
Since COVID-19 I eat more unhealthy	1			1		
Since COVID-19 I exercise less	1			1		
I am worried about my financial situation	1				1	
The COVID-19 outbreak will make society more unequal	1				1	
I am worried about the health of my children	1	1				
I am worried about my children s education	1	1				
COVID-19 will increase divorce rates	1			1		
I am anxious about not being able to meet with friends	1					1
I worry how living in isolation will affect me	1					1
The government is doing a good job dealing with COVID-19	1	1				
The government discloses real numbers of coronavirus infections and deaths	1				1	
The restrictions caused by COVID-19 will continue at least until the fall	1					1
Middle school	1		1			
income don't want to answer	1			1		
# of children 0	1	1				

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 7
The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience

Elissa Moses
Columbia University, USA & BrainGroup Global, USA

Kimberly Rose Clark
https://orcid.org/0000-0003-4712-0475
Dartmouth College, USA & Bellwether Citizen Response, USA

Norman J. Jacknis
Columbia University, USA & Intelligent Community Forum, USA

ABSTRACT

This chapter summarizes the role that artificial intelligence and machine learning (AI/ML) are expected to play at every stage of advertising development, assessment, and execution. Together with advances in neuroscience for measuring attention, cognitive processing, emotional response, and memory, AI/ML have advanced to a point where analytics can be used to identify variables that drive more effective advertising and predict enhanced performance. In addition, the cost of computation has declined, making platforms to apply these tools much less expensive and within reach. The authors then offer recommendations for 1) understanding the clients/customers and users of the products and services that will be advertised, 2) aiding creativity in the process of designing advertisements, 3) testing the impact of advertisements, and 4) identifying the optimum placement of advertisements.

DOI: 10.4018/978-1-7998-6985-6.ch007

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

This chapter explores the resources and climate for creating more impactful advertising through the potential use of Artificial Intelligence (AI) and Machine Learning (ML) in three parts:

1. **The Vision:** Why is AI/ML important to advertisers? What might it be able to deliver now and, in the future, given what is known about the impact of advertising on brain and behavior?
2. **The Theory:** What evidence do we have concerning the key variables that we can use to inform and optimize our advertising (measurable image and sales driving) and, in turn, predict performance? How are academics and marketers creating data bases and conducting meta-analyses to determine the most powerful drivers for implementing AI/ML?
3. **The 'Know How':** What are the tools and methods of AI/ML that show the greatest potential for identifying drivers and predictors of advertising excellence?

MAIN FOCUS OF THE CHAPTER

The Vision

Since the golden age of advertising in the 1960's through the 2000's, before rapid paced digital ads became communications currency, it has been held sacred that *"advertising is an art, never a science."* There has long been the romantic notion of the copywriter facing the *tabula rasa* or blank page, perhaps brimming with ideas or tormented to find the right balance of art, entertainment and salesmanship. This paved the way to high salaries, prestige and awards at Cannes when inspiration and craft hit the bullseye to work advertising magic.

This conventional wisdom has recently been challenged as several factors collided to invite change:

1. **Neuroscience and Behavioral Science** have developed better practical and affordable methods for measuring critical processes that drive consumer behavior such as Attention, Cognitive Processing, Emotional Response and Memory Encoding.
2. **Key Performance Indicators (KPIs)** are now accessible within large data bases that have been amassed. These KPIs can now be linked to effective advertising within brands and categories, particularly nonconscious emotional factors, giving us clues about the characteristics that produce effective advertising in market.
3. **Artificial Intelligence and Machine Learning (AI/ML)** have advanced to a point where analytics can be used to identify variables that drive more effective advertising and predict enhanced performance. In addition, the cost of computation has declined making platforms to apply these tools much less expensive and within reach.

Insights and discoveries by academics and market researchers have identified the critical drivers of effective advertising and software has been developed to crank out "acceptable" digital ads. Is it only a matter of time before advertising will be routinely created, at least in part, by algorithm? We are rapidly developing heuristics and principles for best practices and for advanced evaluation and prediction of efficacy.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

We should expect, even embrace, the development of all kinds of advertising using AI/ML as a starting point for best practices harnessed by human creativity. This will allow us to develop strategies that will catapult the creative process to a higher level of effectiveness and ultimately impact positive return on investment (ROI).

In essence, we are not suggesting that the industry abandon creatives, but instead that we do away with the unnecessary blank page. By embracing AI/ML applications in advertising creation, we can leap ahead to better strategies and themes as a more productive starting point, making the creative's "hit rate" more successful. In this sense, AI/ML has the potential to strengthen the creatives' job security rather than threaten it, as some might fear.

What needs to be recognized, sooner than later, is that technology always has a way of forging ahead and touching every category where it can be made useful. Those who ignore the promise of AI/ML for making advertising and marketing communications more effective, potentially faster, and less expensive, are likely to become less relevant within their own fields. In contrast, those who embrace the opportunity will become part of the conversation and likely industry leaders and experts for the new age.

There are many misconceptions about how advertising works, with conventional wisdom placing unrealistic emphasis on ad recall and conscious brand purchase intention. Recent advances in neuroscience have put us leap years ahead in our understanding of how the brain processes information, stimulates emotional response and utilizes nonconscious associations. Unless there is a strong call to action (such as a timely product promotion, an immediate need for a product, breakthrough messaging, or high entertainment value) ads will largely continue to have less of an impact on conscious brand perceptions and behavior. This is because the average person sees up to 10,000 advertisements every day (Carr, 2021). How many will the average person remember on any given day?

The majority of ads create nonconscious associations and perceptions about a brand that then live in a consumer's subconscious - triggered only when the category or brand becomes salient for consideration and choice, activated in discussion, or referenced in some other way. If asked to describe how one feels about a brand, a confluence of factors ranging from product experience, brand reputation, social media presence, word of mouth, and overt advertising often become entwined to form a single gestalt conscious impression. We also know from neuroscience that emotion is a dominant influencer of our perceptions. Hence, if advertising creates positive brand associations, as it often does in being the hero in a problem solution situation or positive experience depiction, the emotions that are tied to these associations have a residual effect when the critical purchase consideration occurs. In contrast, positive emotions that are tied to associations for a competitive brand, or negative brand associations driven by negative word of mouth or a poor product experience, can be equally impactful in deterring selection.

This describes a macro view of how advertising works from a brain processing perspective. But what if we were to reverse engineer ad effectiveness from an empirical and elemental analytics perspective, using AI and ML to determine the elements that are most critical for success?

To date, multiple studies have been conducted using various neurometrics to determine the key characteristics of "high achieving" ads. Most of these metrics converge on key findings such as gaining attention in the first few seconds, provoking intense emotion (be it positive or negative) to tell an engaging and compelling story, and/or evoking positive emotion and attention around key branding moments. These elements are highly predictive of ads that will do well in the market (the best indicator when data is available), win awards and score high on rational impact.

In 2017, a pivotable study was supported by the Mars Company (USA) in collaboration with multiple market research agencies (MediaScience (USA), Realeyes (UK), etc.) and published in the *Journal of*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Advertising. Through this work, Mars commendably provided a data base of 100 ads across numerous categories that were classified by in-market experience in terms of their degree of success against sales objectives.

The mining of data from facial action coding (FAC; Realeyes (UK) and Professor Maja Pantic), electrodermal activity (EDA), and heart rate variability (HRV) (MediaScience (USA) and Dr. Duane Varan) has revealed critical insights:

1. **Neurometrics** predict in-market success. This is based on the high correlation observed between these relevant neurometrics and corresponding in-market success criteria.
2. **Other critical measures** of interest to advertising presentations have been identified as critical correlates and drivers of in-market performance including attention, emotional arousal, positive affect (at critical moments such as brand identification) and head pitch (an indicator of interest).

The identification of these drivers means that advertisers can deliberately measure these performance milestones to predict a greater probability of ad success. Moreover, the application of AI/ML to delve deeper into the critical elements of effective ad components can reveal additional invaluable insights into getting ad development right from the start.

The Theory

There have been numerous academic studies that have paved the way for AI/ML to make a strong positive impact on the industry, particularly with regard to predictive analytics for modeling, thereby influencing favorable outcomes. Below we examine some of these pivotal studies and outcomes that are making an important contribution to what is possible.

Numerous academics and practitioners are exploring the best ways to harness data to create better predictive advertising models. As the number of available implicit measures continues to grow in terms of their utility, ease of use, and efficacy, algorithmic prediction models will be created and implemented to allow for multiple potential inputs; from neural psychophysiological data streams to self-reported responses. From there, practical workflows will be devised to integrate any number of data stream inputs. Such models will accommodate multiple data inputs before, during, and after a branded experience. Once collected, traditional statistical models or ML models could be used to resolve issues of covariance among these data streams to best explain or predict consumer behaviors.

Such a model would allow for the creation of 'ground truths' among these data streams. One such model is referred to as System N (Clark, K.R. Moses, E., Leslie, K.R., In prep). This model takes into consideration the fact that Dual-process theories of consumer decision making, comprising both intuitive (System 1) and trained (System 2) processes, dominate research in consumer neuroscience. System N provides a holistic framework for integrating explicit (verbal) responses with the diverse array of implicit brain and body measures now available (e.g., electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), heart rate (HR), EDA, automated FAC, etc.). System N is based on our current understanding of relevant contextual 'priors' that shape consumer response (e.g., attention, emotion, habit, personality traits, etc.). This framework provides insights for understanding the dynamics of decision making to better train support vector models (SVMs) of ML algorithms in order to better predict consumer behavior at the market level.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Using fMRI to NeuroForecast Market-Level Outcomes

Historically, behavioral prediction models of market-level choice have outperformed neural models. More recently, however, it was found that neural responses can predict market-level choice and outperform behavioral measures. This seminal research was conducted within a novel internet crowdfunding context (Genevsky et al., 2017). Using both targeted and model-free analyses, it was found that nucleus accumbens (NAcc) activity can predict internet crowdfunding. To elaborate, researchers hypothesized that neural activity could forecast market-level crowdfunding outcomes weeks later. In this study, participants decided whether to fund potential projects that were described on an online crowdfunding website while changes in localized brain activity was recorded via functional magnetic resonance imaging (fMRI). Activity in both the NAcc and medial prefrontal cortex (mPFC) predicted whether individuals chose to fund a potential project. However, only NAcc activity predicted market funding outcomes that occurred weeks after the study. Behavioral measures, on the other hand, did not forecast these market funding outcomes. Thus, brain-based predictors of individual choice can be accurate predictors of market-level crowdfunding outcomes. In addition to providing novel evidence for neuropsychological processes implicated in crowdfunding choices, these findings highlight the usefulness of neural data for predicting choice at the group level – findings that could one day inform applications related to purchase behavior or policy acceptance.

Below are examples of metrics that academic research has shown to be relevant data sources for improving prediction models with AI/ML:

1. **Eye Tracking and ML:** Using support vector machines to classify the first 15 seconds of eye tracking data, Pfeiffer et al. (2020) demonstrated similar levels of correct classifications of search motives when participants viewed supermarket shelves in virtual reality (80%) or physical reality (85%). Applying an ensemble method increased the prediction accuracy to approximately 90%. These findings provide evidence that in virtual reality, information search behavior is similar to that used in physical reality. This finding opens up a whole new realm of possibility with respect to using AI/ML to model VR for testing and predicting consumer response outcomes.
2. **Eye Tracking and FAC:** In this novel means of emotion classification, Gill & Singh (2020) provide a useful review of the numerous neuromarketing techniques currently available. These include integrated automated eye tracking and FAC for emotion analysis to measure arousal and valence responses to various product presentations. In their Table 2, they provide a very useful spreadsheet comparing these techniques in terms of their recommended usage as well as their pros and cons. Techniques covered include fMRI, EEG, eye tracking, pupillometry, FAC, additional biometrics (e.g., EDA, HR, etc.) and additional emotional measurement techniques (e.g., sentiment analysis).
3. **EEG and ML:** Hakim et al (2020) discuss their use of multiple types of measures extracted from EEG recordings and ML algorithms to improve preference prediction above using self-reports alone. EEG was measured as subjects watched video commercials of six food products and then they responded to a questionnaire. Next, subjects made binary preference choices of the food products. The researchers attempted to predict within-sample and population level preferences. Using these methods, researchers were able to predict subjects' most and least preferred products with 68.5% accuracy, an improvement in accuracy of four percentage points compared to predictions that were based on self-reports alone. In addition, EEG measures improved within-sample prediction of all six products by 20%. Moreover, at the population level, EEG measures increased prediction by

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

12% compared to questionnaire-based predictions. The most predictive EEG measures were frontal powers in the alpha band, hemispheric asymmetry in the beta band, and inter-subject correlation in delta and alpha bands. Hence, by employing multiple types of EEG measures and ML models, Hakim et al (2020) showed that they could predict consumers' preferences and the success of marketing campaigns at the population-level.

4. **Fusion of EEG Response and Sentiment Analysis:** Kumar et al. (2019) have proposed a novel multimodal framework to predict consumer ratings of presented products by fusing different data sources, namely physiological signals with product reviews. Product reviews that were posted by viewers were retrieved using a Natural Language Processing (NLP) technique to compute a global product rating. Also, participants' EEG signals were recorded while they watched different products. Subjects then provided a valence score where a higher valence score indicated a higher intrinsic attractiveness of a product. Random forest-based regression techniques were used to model the EEG data to build a rating prediction framework. Further, an Artificial Bee Colony (ABC) based optimization algorithm was used to boost the overall performance of the framework by fusing global and local ratings. EEG from 40 participants (25 male) was recorded while they viewed 42 different products. Results suggested that the proposed ABC optimization approach can achieve lower Root Mean Square Error (RMSE) in rating predictions compared to individual unimodal schemes.
5. **One Person fNIRS:** Functional near-infrared spectroscopy (fNIRS) is a noninvasive optical imaging technique that measures changes in hemoglobin (Hb) concentrations within the brain via the absorption spectra of Hb in the near-infrared range. fNIRS provides an alternative approach to the acquisition of task-related hemodynamic signals using optical methods with the advantage of allowing for data collection in more natural, upright, and face-to-face conditions (compared to fMRI, for example).
6. **Two-Person Neuroscience:** Most of what we know about the social brain is based on fMRI studies that involve static and non-interactive conditions. Thus, data assessing the neural mechanisms of live "in-person" social interactions between subject dyads is sparse. Recent work addresses the advantages of using fNIRS to study these live two-person social interactions. These studies are often conducted in the context of the Interactive Brain Hypothesis which proposes that interpersonal interaction between individuals evokes neural mechanisms not engaged during solo, non-interactive, behaviors (Hirsch 2020; Hirsch et al, 2017; Holper et al, 2012). In one study (Hirsch et al, 2017), subject dyads who knew each other and those who were strangers were recruited. Then, fNIRS activity was recorded while dyads engaged in direct eye-to-eye contact (interactive condition) or a condition where subjects engaged in mutual eye contact with a photograph of a face (non-interactive condition). Signals acquired during the interactive condition identified a 'within-brain' left frontal network that included pars opercularis, pre- and supplementary motor cortices, the left superior temporal gyrus, and primary somatosensory cortex. In addition, this condition identified functional synchrony (i.e., 'cross-brain' coherence) within left superior temporal, middle temporal, and supramarginal gyri, as well as the pre- and supplementary motor cortices of both interacting brains. Interestingly, this synchronous cross-brain activity was partner-specific (i.e., disappeared with randomly assigned partners). These exciting findings demonstrate the power of fNIRS to provide a more ecologically valid means for the study of interpersonal interactions.
7. **A Synchrony-Based Classification Approach for Predicting Attitudes Using fNIRS:** People who share similar viewpoints show greater neural synchrony (i.e. correlated neural fluctuations over time) with one another. These effects have been demonstrated at the group level. Dieffenbach

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

et al. (2021) used a simple synchrony-based classification method, referred to as the 'neural reference groups' approach, to predict individuals' dispositional attitudes. Using fNIRS, in a mobile 'pop-up' neuroscience lab, Dieffenbach et al. (2021) compared the amount of synchrony subjects showed with the brain activity of two partisan neural reference groups. Results showed that the authors could significantly predict individuals' partisan stances on a sociopolitical issue based upon their degree of synchrony with one group versus the other. They found that dorsomedial prefrontal cortex (dmPFC) activity specifically predicted an individual's partisan stance at above-chance levels. Thus, this 'neural reference groups' approach can be used to investigate naturally occurring sociopolitical opinions anywhere in the world.

8. **Color as a Context:** There are also studies that identify other key variables that are highly relevant for predicting strong ad performance or contributing to the creation of effective advertising. Noting those variables and including them into the modeling mix is expected to be fruitful for shaping and influencing advertising outcomes. Bazley et al. (2021) have shown that using the color red to represent financial data impacts individuals' risk preferences, expectations of future stock returns, and trading decisions. Interestingly, these effects were not observed in individuals who are color-blind. Further, the effects were not as robust in China (where red represents prosperity). Hence, cultural context significantly impacts this effect. Similar effects were not observed for other colors, including yellow and blue. But again, it could be the case that the color red is only relevant in this context. In essence, when certain colors are featured, simple perceptual processes are shown to influence complex behavior. In this case the effect was measured upon investor behavior. In other circumstances, such as soft drinks, the color red might be expected to induce thirst and preference for Coca-Cola.

The "Know How"

There are numerous tools and methods of AI/ML that are being used by academics and industry practitioners that have a high potential for identifying drivers and predictors of advertising excellence.

As is true for almost every other activity, people are finding increasingly valuable ways to use AI/ML to augment their own effectiveness and skill in the practice of their professions. However, the sheer volume of data about consumer markets makes this research especially appropriate for the AI/ML methods that exist due to "big data."

Indeed, one of the aspects that distinguishes modern ML from traditional statistics is the focus on what can be done with large amounts of data. Historically, statistical methods were focused more on a paucity of data (as a result of the expense of collecting it) and whether such a data sample accurately represented some truth about the much larger population from which it was drawn.

For reference, listed here are some of the key methods employed in AI/ML with their potential application for advertising and marketing:

1. **Neural Networks:** These models are referred to as Deep Learning and are often what people think of when they use the term artificial intelligence in general. Neural networks (particularly multi-layer networks) create models that (usually, but not always) provide the best predictions and classifications; especially for images, video and other non-traditional data. These models are powerfully applied to the optimization and creation of advertising by predicting desired outcomes based upon visual features. However, their accuracy comes with a price. Neural networks are often difficult to

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

explain and understand and require much more computing power than other methods, especially other methods of supervised learning.

2. **Regression Analysis:** Regression analysis is used to determine which variables are having the most impact on a certain outcome. It is based upon linear relationships between continuous variables that can be used to predict a value, like the number of sales or the number of errors. It is a well-established analytic tool (before AI/ML became household words) for understanding the critical drivers in a data set. The predictors (called independent variables) should not be strongly correlated and should not be dependent on each other.
3. **Decision Tree Methods:** Decision trees are among the most powerful and popular tools for classification and prediction. The method is ideal for identifying group membership, especially target groups for advertising and marketing. These methods classify people (or things) to predict what group they would be in and is a very common use of ML. The results from decision tree methods are easy to explain, having the advantage of being able to identify non-linear combinations of variables or profiles to determine the probabilities for being in various groups.
4. **Clustering:** Clustering lets the machine determine how different variables group together without any guided indication as to what classification is right or wrong. Hence it is an example of 'unsupervised learning' and has many applications, such as reducing the number of variables in a dataset or identifying groups with correlated characteristics. Although the clusters are not always clearly explained and must be interpreted for meaning and application, they have been used widely in marketing – for example, psycho-demographic categories like *'soccer moms'* in the US.
5. **Network Analysis:** It is important to understand the impact of one-person, physical object, or variable on others in a connected world. This is different from most other methods that deal with a large number of observations about people or objects. Network analysis enables the identification of complex patterns or relationships with a single individual, object or variable and has applications for understanding social networks, media consumption patterns, word-of-mouth marketing and the like.
6. **Natural Language Processing and Text Analytics:** Much data is in the form of text and easy to input for analysis. Though content analysis has existed for many years, the more recent use of chatbots and language translators is a reflection of major advances in natural language processing by machines. Text analytics can extract key concepts (Content Analysis) and assess the sentiment of the text (Sentiment Analysis); both of which are useful in understanding consumer reactions to products and their advertising. More recent advances have created software such as the Generative Pre-trained Transformer 3 (GPT-3) which is an autoregressive language model that uses deep learning to produce human-like text.

SOLUTIONS AND RECOMMENDATIONS

Along with the enormous amount of data that are now available, these AI/ML tools can play a significant role in enhancing each of the major phases of advertising:

1. **Understanding** the clients/customers and users of the products and services that will be advertised
2. **Aiding** creativity in the process of designing advertisements
3. **Testing** the impact of the advertisements

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. **Identifying** the optimum placement of the advertisements

Understanding Clients/Customers and Users

As always, in order to develop effective advertising, creative developers need to deeply understand the target of the advertising – the clients, customers, users whose decisions and behavior will determine success or failure. There is now a plethora of data about most consumers that enable very nuanced views of consumer segments.

Customarily, in thinking about segments and user personas, marketers have used various general psychographics and terms like 'metrosexual' (i.e., a man who is sensitive and fashion-conscious), coined by trends guru Marian Salzman, or 'soccer moms' (i.e., suburban, likeminded moms) to target specific demographic groups. With ML, new models of segmentation can be created that are specifically intended to aid in understanding a market for a particular product with greater distinction and granular nuance.

Methodologically, the various forms of *cluster analysis* are ideal for this purpose, whether they be the popular *k-Means method* or methods that share some of the characteristics of *neural network models*, like *Kohonen group mapping*. These methods identify clusters or segments of the market that can help form a holistic sense of the various kinds of individuals to target.

These clustering methods have also been used in medical and manufacturing applications to identify anomalies — smaller outlying clusters of cases that highlight problems from cancer cells to defective machinery. In advertising, however, these smaller clusters may help identify promising sub-segments of the market that can be reached with customized messaging and that competitors may be ignoring.

Aiding the Design Process

Perhaps the area of greatest controversy, but greatest potential, is in the use of AI to aid and inform the creative design process. This is, in particular, a reflection of the advancements in what is commonly called *"deep learning,"* the use of many-layered, complex *neural networks* and the various modifications to neural network models that are called *generative adversarial networks* (GANs).

1. Identifying Features of Ads

Advertising professionals have been taught to look at a select, finite set of features in crafting ads. While this perceived wisdom cannot be casually dismissed, there may be other aspects of ads that must be considered, especially those that are more nuanced or are at a more granular level than people consciously notice.

Deep learning can identify these aspects. In the ML world, these are called *extraction of features*. These features cover all aspects of the ad, including visual characteristics, audio characteristics, the text of the words used, and its pace and timing. For instance, the example of the inherent information communicated by the color red (noted above) is a clear example.

In an additional case example, Airbnb has used deep learning tools to analyze the impact of pictures of the lodging that its member hosts are renting, largely in an attempt to help the hosts present more attractive views and to generate more revenue for both the hosts and Airbnb.

Neural networks underlie the various tools that identify the emotional reactions of consumers while they watch advertisements or view the product itself. Among the leaders in this kind of emotional or

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

affective FAC analysis are Realeyes (UK) and Affectiva (USA). Another approach, using the work pioneered by audEERING (Germany) is to apply voice analytics in order to measure the emotional content and tone of speech in advertising.

If a feature or distinct element of advertising is present and seems to play a role in ad effectiveness, *deep learning* can identify it. For example, a more bluish-hue in the background may end up being more effective than a more reddish-hue in the background. Going into the analysis, the analyst may have an instinct about this or have never really thought about it. In either case, the *deep learning* model will help to identify which of the existing elements are most informative for what kinds of situations and what new elements and ideas might be added to increase the effectiveness of the ad.

Finally, a word of advice: AL/ML can take you to unforeseen discoveries in data analytics that leapfrog linear logic and enable you to see drivers, patterns, relationships, and insights that were never before possible with traditional methods. However, the key is that the attitude of a person embarking upon ML needs to be open. This means, in essence, connecting with your inner five-year-old, who asks lots of questions, hoping that the model will answer those questions, even if only in part. Ad professionals can adopt this childlike sense of possibility or ignore it at their own peril, as newcomers to the field will advance by looking at the world through unassuming, nonjudgmental eyes.

2. Assessing the Role of These Features in Context

Ultimately, the big question for those in marketing is *"What leads to purchase?"* Consider ML models aimed at predicting sales. Clearly, there are non-marketing factors that affect sales – the overall economy, the objective quality of the product, or pricing versus the competition. And yet, it is possible to add advertising into the mix and effectively isolate its impact. Moreover, the possibilities go beyond our usual attempts to simply see what impact ad spending has had on sales.

With the measurement of the many ad features that have been identified, we can add a variety of new variables to the model, including the hue of the background, the likeability of the music, the appetite appeal of the food shown, or even the presence of dogs, brides and babies (which conventional wisdom suggests are elements that capture attention). These variables may or may not turn out to be important, but that's the point of doing the analysis. As more data are collected and more analyses are done, professionals will learn that some of what was believed to be important turns out to be less so; and some of what was ignored turns out to be more so. Data scientists may call it ML, but any field that uses ML soon realizes that the humans are also learning right along with the machines.

3. Enhancing the Creative's Toolkit

Already, there are companies offering AI software that can create advertising copy, jingles, music, and even (albeit not always funny) jokes. In one well publicized move, Chase Bank decided to replace its human copywriters with AI software[1]. In late 2020, the advances of yet another generative approach, GPT-3, opened the possibility of even more flexible creation of advertising by an AI package[2].

However, in thinking about the use of AI, it may help to set technology aside for a moment and go back more than 500 years. Consider oil painting – the "modern" European use of oil-based paints dates back to that time. These paints had different characteristics than the materials that were used earlier. But this new medium, including the mix of oil paints, brushes and easel, gave artists a new range of ways to

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

be creative. Some of Europe's greatest works of art resulted. Similarly, ad designers can think of AI as a new tool that can dramatically expand their creative options.

Although *neural network* models have often been able to provide more accurate predictions of traditional data about numbers than models based on other methods, they have been especially valuable in dealing with non-traditional data, including images, video, and audio.

It is likely, at least for the next several years, that human creativity informed by the results of AI-based measurement of consumer reactions will be the basis for the most successful advertising, rather than an exclusive dependence on either traditional human-only creative advertising work or AI-based software alone.

However, it is not only consumer reactions that AI can identify. *Neural network* models can also extract visual features that distinguish one advertisement from another in ways that designers have not been taught to think about. The features it identifies may not always be of relevance, but it is worthwhile for the creative or designer to consider them. It is also useful to determine which of these features are correlated with positive reactions to the advertisements. With this understanding, it is likely that the human can create more innovative and impactful advertising.

Another application might be in the use of GANs to create what are commonly called 'deep fakes,' images that look like real people doing something that actually never happened. It would be possible for a designer to use a GAN to mesh together pictures of people representing various key population segments into one 'combined' image[3]. That combination would not be clearly identified as a particular individual, yet it would still look very human and could play a necessary representative role of a consumer in an advertisement. Clearly there are ethical and legal issues in using artificial intelligence technology this way. However, this capability of GANs can be used positively in the support of advertisements.

Testing the Impact

Given that the use of ML is part of the human learning process, the analysis does not stop with a happy *"eureka!"* moment based upon the first model that is generated by the computer. That and further iterations will need to be tested and, in this way, the model is improved by additional data. This is not so different from the more general idea that an initial ad may have to be improved to become more effective after testing or the market has made its judgment.

Thus, many of the tools and methods that are suggested for the initial process of developing an advertisement can also be used to test the impact of the advertisement. However, more can be done once the product and the advertisement face the test of the market. Whereas creative teams are limited to seeing the reaction of consumers and perhaps the consumers verbal feedback in the design phase, the market will eventually test this design with sales data.

At this stage, the goal of the ML models would be to establish which characteristics of the advertisement increased sales. That may mean that consumer reactions will shift from being the main target, to being one of a number of factors that are used to predict sales.

For example, in their promotional copy, Affectiva has written that "Over the last decade, advertisers have increasingly focused on creating emotionally resonant ad content, with the knowledge that emotions are a key predictor of sales and brand loyalty"[4]. While this is a reasonable assertion, is not yet clear how much of a "key predictor" emotions are, and indeed the influence is likely to vary depending upon the products, the consumer targets, and the branding strategy of the advertiser. With ML, it is now possible

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to get a much more precise understanding of the relationship between these emotional reactions and ultimate purchase decisions.

It is also worth noting that the traditional cycle of ad creation, followed by ad assessment, and then followed many months or a year later by a new ad, is no longer the norm in a digital world. It is now easier to adjust the ad as a result of a quicker analysis of consumer testing or market reaction. Thus, AI/ML become integral to a continuous process of design, feedback and re-design to achieve optimal results.

Identifying Optimum Placement

Developing the ad is only part of the story, of course. The ad has to be placed somewhere for target consumers to see it. While there are an abundance of data provided by media outlets about their audiences, these data are not necessarily directly tied to the purpose of the advertiser.

In the current world of big data, it is possible for the advertiser to take advantage of superior models to determine which media work best for a given advertisement or segment of the market. This requires a level of granularity in tailoring advertisements to market segments that have not been commonly used, but that can now be more easily identified and targeted with ML. This is one use of *decision trees*, especially ensemble versions that substantially increase accuracy. Moreover, these methods, including ML using network analytics, can also help with pursuing digital social network strategies to get the advertising message out.

FUTURE RESEARCH AND DIRECTIONS

There have been rapid advances in AI/ML in recent years and there is no reason to assume that anything will slow down this advance. Recent ML advances such as real time language translation has gone from a future vision to being 95% comparable to the average human translation. In fact, in market research at HARK Connect, marketers can watch focus groups from anywhere in the world with simultaneous translations and subtitles in over 60 languages.

We are living in a technically advanced world where almost anything imagined can be realized through AI/ML technology. This means that the ability to use AI/ML to advance the effectiveness and even creation of successful advertising will only accelerate as both the models and the humans who control them become ever smarter.

CONCLUSION

We will need a new business model for consultants, strategists, AI/ML experts, creatives, and clients to work together in new ways for superior effect. Pradeep, Appel and Sthanunathan (2019) point out that clients benefit when marketing campaigns are derived from larger reservoirs of information and insights. They also note that agencies benefit when clients have more confidence in the creative solutions offered. But it is more than that. Creatives will benefit by jumpstarting the creative process with inputs and insights from AI/ML that will catapult their inspiration and impact many times over, with record speed for greater branding and behavioral impact.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Consider that Chase Bank did better with AI as their copywriters. What superior quality might be achieved with human copywriters armed with the methods of ML? In the near term, the advertising professionals who will excel will be those who learn how to enhance their skills with AI/ML at every stage of advertising development, assessment, and execution.

REFERENCES

Bazley, W. J., Cronqvist, H., & Mormann, M. (2021). Visual Finance: The Pervasive Effects of Red on Investor Behavior. *Management Science*. https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2020.3747

Carr, S. (2021). *How Many Ads Do We See A Day In 2021?* https://ppcprotect.com/how-many-ads-do-we-see-a-day/

Clark, K. R., Moses, E., & Leslie, K. R. (2021). System N: A Multimodal Approach to Understanding Consumer Decision Making. *Journal of Consumer Psychology*.

Dieffenbach, M. C., Gillespie, G. S., Burns, S. M., McCulloh, I. A., Ames, D. L., Dagher, M. M., Falk, E. B., & Lieberman, M. D. (2021). Neural reference groups: A synchrony-based classification approach for predicting attitudes using fNIRS. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 117–128. doi:10.1093can/nsaa115 PMID:33025001

Genevsky, A., Yoon, C., & Knutson, B. (2017). When Brain Beats Behavior: Neuroforecasting Crowdfunding Outcomes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(36), 8625–8634. doi:10.1523/JNEUROSCI.1633-16.2017 PMID:28821681

Gill, R., & Singh, J. (2020). A Review of Neuromarketing Techniques and Emotion Analysis Classifiers for Visual-Emotion Mining. *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART),* 103–108. 10.1109/SMART50582.2020.9337074

Hakim, A., Klorfeld, S., Sela, T., Friedman, D., Shabat-Simon, M., & Levy, D. J. (2020). Machines learn neuromarketing: Improving preference prediction from self-reports using multiple EEG measures and machine learning. *International Journal of Research in Marketing*. Advance online publication. doi:10.1016/j.ijresmar.2020.10.005

Hirsch, J. (2020). *An fNIRS Approach to Two-person Neuroscience*. Biophotonics Congress: Biomedical Optics 2020 (Translational, Microscopy, OCT, OTS, BRAIN) (2020), Paper BM4C.1, BM4C.1. 10.1364/BRAIN.2020.BM4C.1

Hirsch, J., Zhanga, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage*, *157*, 314–330. doi:10.1016/j.neuroimage.2017.06.018 PMID:28619652

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1), 212–222. doi:10.1016/j.neuroimage.2012.06.028 PMID:22732563

Kumar, S., Yadava, M., & Roy, P. P. (2019). Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Information Fusion*, *52*, 41–52. doi:10.1016/j.inffus.2018.11.001

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pfeiffer, J., Pfeiffer, T., Meißner, M., & Weiß, E. (2020). Eye-Tracking-Based Classification of Information Search Behavior Using Machine Learning: Evidence from Experiments in Physical Shops and Virtual Reality Shopping Environments. *Information Systems Research*, *3*(3), 675–691. doi:10.1287/isre.2019.0907

Pradeep, A. K., Appel, A., & Sthanunathan, S. (2019). *AI for Marketing and Product Innovation*. Wiley.

ADDITIONAL READING

Cheddak, A., Ait Baha, T., El Hajji, M., & Es-Saady, Y. (2021). Towards a support system for brainstorming based content-based information extraction and machine learning. In M. Fakir, M. Baslam, & R. El Ayachi (Eds.), *Business Intelligence. CBI 2021. Lecture Notes in Business Information Processing* (Vol. 416). Springer., doi:10.1007/978-3-030-76508-8_4

Choi, J.-A., & Lim, K. (2020). Identifying machine learning techniques for classification of target advertising. *ICT Express*, *6*(3), 175–180. doi:10.1016/j.icte.2020.04.012

Davenport, T. H. (2019). *The AI Advantage: How to Put the Artificial Intelligence Revolution to Work*. MIT Press.

Gharibshah, Z., Zhu, X., Hainline, A., & Conway, M. (2020). Deep learning for user interest and response prediction in online display advertising. *Data Sci. Eng.*, *5*(1), 12–26. doi:10.100741019-019-00115-y

Jog, S., Siras, B., Fender, A., Mandurkar, P., Nikalje, Y., & Chhabria, S. (2021). Video popularity prediction using machine learning. *International Research Journal of Modernization in Engineering Technology and Science*, *3*, 778-783. https://irjmets.com/rootaccess/forms/uploads/IRJMETS733901.pdf

Lee, J., Jung, O., Lee, Y., Kim, O., & Park, C. (2021). A comparison and interpretation of machine learning algorithm for the prediction of online purchase conversion. *Journal of Theoretical and Applied Electronic Commerce Research*, *16*(5), 1472–1491. doi:10.3390/jtaer16050083

Li, H. (2019). Special Section Introduction: Artificial Intelligence and Advertising. *Journal of Advertising*, *48*(4), 333–337. doi:10.1080/00913367.2019.1654947

Nikolajeva, A., & Teilans, A. (2021). Machine learning technology overview in terms of digital marketing and personalization. *Communications of the ECMS*, 35. http://www.scs-europe.net/dlib/2021/ecms2021acceptedpapers/0125_ocms_ecms2021_0052.pdf

KEY TERMS AND DEFINITIONS

Artificial Intelligence: The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cognitive Neuroscience: The scientific field that is concerned with the study of the biological processes and aspects that underlie cognition, with a specific focus on the neural connections in the brain which are involved in mental processes.

Consumer Neuroscience: Is the combination of consumer research with modern neuroscience. The goal of the field is to find neural explanations for consumer behaviors. Unlike traditional market measures which rely on conscious recall of past experience or deliberative beliefs regarding future behaviors, consumer neuroscience relies on in-moment measures of central and peripheral nervous system processes that often occur outside of conscious awareness.

Deep Learning: A subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that is unstructured or unlabeled. Also known as deep neural learning or deep neural network.

Electrodermal Activity (EDA): A psychophysiological measurement of skin moisture (i.e., sweat) that enhances electrical conductance across the skin. EDA can be measured in as a response to either external or internal stimuli and is believed to reflect how physiologically arousing the stimulus is perceived to be.

Electroencephalography (EEG): EEG is a test used to evaluate the electrical activity in the brain. An EEG tracks and records brain wave patterns that are highly time locked to a stimulus.

Emotion: A natural instinctive state of mind deriving from one's circumstances, mood, or relationships with others.

Facial Action Unit Coding (FAC): FAC refers to a set of facial muscle movements that correspond to a displayed emotion.

Functional Magnetic Resonance Imaging (fMRI): Functional magnetic resonance imaging or functional MRI (fMRI) measures brain activity by detecting changes associated with blood flow. This technique relies on the fact that cerebral blood flow and neuronal activation are coupled. When an area of the brain is in use, blood flow to that region also increases.

Functional Near-Infrared Spectroscopy (fNIRS): Is a non-invasive imaging method involving the quantification of chromophore concentration resolved from the measurement of near infrared (NIR) light attenuation or temporal or phasic changes.

Machine Learning: Is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

Sentiment Analysis: The process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine positive, negative, or neutral attitudes associated with test element.

Supervised Learning (SL): The machine learning task of learning a function that maps an input to an output based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples.

Unsupervised Learning: A machine learning technique in which the users do not need to supervise the model. Instead, it allows the model to work on its own to discover patterns and information that was previously undetected. It mainly deals with the unlabeled data.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

1 Adrianne Pasquarelli (2019). Chase commits to AI after machines outperform humans in copywriting trials. AdAge, July 30, 2019. Retrieved from https://adage.com/article/cmo-strategy/chase-commits-ai-after-machines-outperform-humans-copywriting-trials/2187606

2 Cade Metz (2020). When A.I. Falls in Love. New York Times, Nov. 24, 2020. Retrieved from https://www.nytimes.com/2020/11/24/science/artificial-intelligence-gpt3-writing-love.html. Also see Arram Sabeti (2020). GPT-3: An AI that's eerily good at writing almost anything. Retrieved from https://arr.am/2020/07/09/gpt-3-an-ai-thats-eerily-good-at-writing-almost-anything/

3 For example, see Bill Schmarzo (2019). How GANs and Adaptive Content Will Change Learning, Entertainment and More. Data Science Central, September 8, 2019. Retrieved from https://www.datasciencecentral.com/profiles/blogs/how-gans-and-adaptive-content-will-change-learning-entertainment

4 Hailey Melamut (2020). Research Shows Advertising is Having a More Emotional Impact on Consumers, But Not All Advertisers Are Using Emotions Effectively. Business Wire, October 27, 2020. Retrieved from https://www.businesswire.com/news/home/20201027005291/en/Research-Shows-Advertising-is-Having-a-More-Emotional-Impact-on-Consumers-But-Not-All-Advertisers-Are-Using-Emotions-Effectively

5 Pradeep, A.K., Appel, Andrew, Sthanunathan, Stan (2019) AI for Marketing and Product Innnovation; Wiley, Hoboken, NJ; p 209

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 2
Action

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 8
Retail Customer Native Baskets Creation:
A Must-Do Into Measuring Price Indices

Liliana Bernardino Matias
Sonae MC, Portugal

Ana Costa Freitas
Sonae MC, Portugal

Filipe Fernandes Miranda
Sonae MC, Portugal

ABSTRACT

Standard approaches to measure price index on a specific retailer against competitors tend to be overly optimistic. The authors propose a methodology that deflects from analyzing solely the price of all products effectively bought by customers by taking into consideration their prior shopping intention. Loyalty schemes are game changers when measuring price index. The study proposes a two-step methodology that comprises on building the customer native baskets and allocating the prices collected to each customer. Under the assumption that a price sensitive customer will not buy a poorly priced product compared to the competitors in that specific period, the authors present an innovative methodology resulting in a more realistic price index metric.

INTRODUCTION

Pricing is one of the main elements that influence a customers' purchasing decision, as stated by Hustić et al. (2015), It is essential for retailers to measure their price competitiveness against other players in the market. A common metric used in price comparison is Price Index which can be measured at a stock-keeping unit (SKU) level, for a more granular approach, or at an aggregate level for an overview approach.

DOI: 10.4018/978-1-7998-6985-6.ch008

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The authors claim that measuring price index on a SKU level and grouping it on a broader granularity supported by weights based on effective sales is a fallacy, since the sales themselves are affected by the price index at the SKU granularity.

This chapter aims to introduce an innovative framework for measuring price index in retail. The authors also propose that personalizing communication to consumers should take into account their own price indexes and not overall price indexes, since different retailers may focus on competitive prices in specific product ranges or target segments.

To avoid measuring price indexes on effective sales, the authors propose to measure the price index on the list of intentions based solely on the customer needs, referred as native baskets. This chapter starts by exploring different approaches to create native baskets per customer that achieve a satisfying balance between uniqueness among baskets and coverage of tickets.

The authors then present the framework on which the price index on a SKU level is distributed by the native baskets and how those are grouped into a customer price index to measure the competitiveness on price of a store. Measuring price index on a customer level involves not only taking into consideration the SKUs on the customers' native baskets, but also defining how to weight from SKU to a native basket level. A tailor-made combination of segments and tickets characteristics to weight SKUs relevant per native basket on a customized view and adapting the competitor stores forces to capture the influence of each competitor price index into a unique price index are some of the challenges overcome in this chapter.

Two examples of retail practical tools are presented as outcomes from getting personalized price indexes: a business tool to follow and measure price indexes on different types of customers and transactions and a customer segmentation based on customer price index for personalized targeting and communication.

Throughout the experiment, the authors faced issues regarding sparsity in transactional data and heterogeneity between the resultant baskets, presenting methodologies to define product granularity and find the right combination of tickets' characteristics as starting point.

Business decisions presented in this chapter were adapted to the highly promotional activity in the retail arena in Portugal.

BACKGROUND

Retail is under very high promotional pressure and personalization is key to survive in such conditions.

Measuring the price index against competitors is a powerful tool on mass marketing to promote products with price continuously below the market or to manage customer perception on a global scale.

The first step into personalizing price on a customer level is to accurately measure it.

Price Index is a weighted average of price in relation to a competitor or a set of competitors for a given product or product group on a given time period (Galkin, 2019).

The process to collect price index involves:

Step 1: Matching products that overlap between the focus retailer and the top competitors

Step 2: Picking accurate prices on products in competitors of interest (updated over time)

Step 3: Compute the index between the retailer price and each of the competitors by product

On a product level, it is quite simple to accurately measure price indexes among retailers. However, combining the price index of a category or overall level is increasingly challenging.

Moving from a product level to a product hierarchy level involves weighting the price index by a business acceptable metric e.g. sales per product, volumes sold, frequency, number of customers covered.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Despite the metric selected all the above are limited to effective customer behavior and do not consider that the current price might be a factor taken into consideration prior to the purchase, and thus directly affecting the price index metric.

The progressively high promotional activity experienced in retail stores (Jean & Yazdanifard, 2015), resulting in products being consistently in and out of promotion, combined with the easy access to tools that enable customers to compare prices between retailers, accentuates the cannibalization both between competitors and/or between promotional periods on the same retailer (Kumar, 1988).

These fluctuations in product final prices make the combined price index overestimation hard to avoid when weighting the product prices by the units sold.

In this chapter, we propose to measure the price index on the customers' list of intentions. These lists are defined as the native baskets and the products at the stock-keeping unit (SKU) level will be weighted on a customer level, resulting in a fully personalized approach.

Under the digital era, creating shopping lists is a scientific area of increasing interest (Yu et al., 2016).

Data mining approaches on sparse datasets such as retail purchases tend to leave out a relevant part of all products. Prassas et al. (2001) propose a recommendation model combining choosing the product category based on association rules followed by collaborative filtering (CF) in order to identify the specific product to be recommended. Implicit ratings and preferences are extracted from the loyalty transactional data as input for the collaborative filtering. Despite being good at tackling sparse transactional data, the limitations are in how to adapt from a product recommendation to a full basket recommendation. Besides that, CF tends to capture customers' general interests but has difficulty in considering sequential features of historical transactions. In addition, it is under the assumption that similar customers share similar product preferences.

Shopping list reminders are often described as basket prediction and intend to cover both the recurrence of purchasing patterns over several weeks and adaptation to consumer behavior evolution over time. To tackle both objectives hybrid approaches are becoming popular.

Doan et al. (2018) propose using generative adversarial networks (GAN) to generate sequences of baskets followed by long short-term memory (LSTM) to model the behavior evolution of a customer. The main downside of artificial recurrent neural networks such as LSTM is the lack of interpretability of their predictive models by humans. Poor performance in terms of time consumption is also an issue to take into consideration when analyzing all transactions of the entire retail loyalty scheme database.

Guidotti et al. (2019) have a user-centric approach that consists of using only individual data through temporal annotated recurring sequences (TARS), capturing the customer's purchasing habits by sequential recurring patterns followed by the prediction of the next basket of the customer through TARS Based Predictor (TBS). Such approach is expected to be under a strong cold start.

Using predictive analysis for a basket prediction has the downside of being highly influenced by the effective historic purchases, the seasonal dependency, the high individuality of the baskets (no homogeneity), and the tendency not to capture the overall behavior of customers.

Creating native baskets is intimately linked to the study of market basket analysis.

Market basket analysis is based on association rules. The frequent item sets (FIS) methodology allows to identify a set of products with a higher presence per customer/basket.

FIS allows to find a set of product groups frequently bought together. Apriori and eclat variants are plausible candidates for the study. Apriori is a widely known algorithm that uses breadth-first search to identify general trends on the dataset. It tends to be very time consuming since it scans the database

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

several times (Wang, & Karypis, 2004). Eclat performs a depth-first search strategy of a prefix tree to formulate association rules, achieving satisfactory running-time performances (Mahanti & Alhajj, 2005).

Limitations regarding FIS methodology is the tendency to generate a high number of patterns with very low length, most of which are spurious and not causally related and thus not relevant to recommendations (Wang et al., 2014).

Clustering tickets is an alternative approach to group similar baskets in terms of its composition and variety, allowing to extract the set of core products on the clusters. This methodology considers all transactions with its whole components.

Clustering transactional data relies on relevant items and irrelevant rare items. Most clustering algorithms struggle facing the large presence of outlier items that do not contribute to the clustering formation (Bouguessa, 2011).

Transactional data can be interpreted as consisting of special categorical data, which can be easily transformed to the traditional row by column table with Boolean values (Yan, Chen, & Liu, 2006). Most clustering approaches adopt a pairwise similarity such as the Dice and Jaccard coefficients for measuring the similarity of two transactions (Wang et al., 1999).

Defining the optimal number of clusters as an input parameter is a big challenge when dealing with real datasets. Iterative test and learn to fine-tune the selected k clusters may become a time-consuming task and sampling the typically large transactional datasets on the assessment step may accelerate the process.

Guidotti et al. (2017) propose a parameter free clustering algorithm to create personal purchasing lists using txmeans algorithms. The outcome is representative clusters but its restriction on a customer-by-customer approach creating baskets has performance issues when escalating to a wide loyalty scheme and generates individualized baskets.

CUSTOMER PRICE INDEX REVIEW

Linking the Main Stages Under Study

Defining a customer price index framework comprises two main stages to be tackled on different depths of detail.

Thus, the authors partition the study in two main stages:

Phase One: Create native baskets
Phase Two: Allocate price to native baskets per customer.

Native Baskets: Methodology Overview

One of the main assumptions regarding native baskets is the heterogeneity among each other and coverage of the overall tickets in store.

One must keep in mind that a customer might have different shopping lists/native baskets depending on their shopping mission.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Predictive Algorithms

Predicting shopping lists aims to predict future purchases taking into consideration individual historic purchases, meaning that the prediction is made on a customer level.

The idea behind native baskets is that it represents the shopping list of a transaction and should be valid across several transactions and customers with the same purpose. By predicting the purchase of a customer, the model will not generate a generalized shopping list, but a personalized prediction. The customer level becomes a limitation to this study, and an approach that looks at groups of customers/ purchasing moments is more desirable.

Serendipity is one of the key elements recommending a list to a customer and it was tackled using e.g. collaborative filtering approaches, but it is highly undesirable under the purpose of the current study.

Market Basket Analysis

Market basket analysis using a FIS methodology easily reaches an outcome with product groups bought together, being a good definition of native baskets.

The good performance of such algorithms is a must, and the high variety of products on the tickets combined with the Apriori breadth-first non-efficient approach turns it into a non-desirable approach for creating native baskets.

The authors recommend never to lose focus from the support of the FIS at the expense of getting niche baskets. The Eclat algorithm is appropriate to confront such issues.

As limitations, the FIS methodology tends to generate a high number of patterns with very small length, thus not contributing with relevant results to the construction of native baskets on the business view. The constraints related to its high restriction to top frequent items may have generic baskets as outcome. Being capable of detecting large patterns for e.g. stock up, native baskets is mandatory for the success of the implementation.

Clustering Analysis

Clustering tickets cover all the transactional information being fed to the model and consider all combinations of items to build the native baskets. As an outcome, if a basket is sufficiently relevant in terms of frequency it is going to end up as a native basket, despite being a niche one.

The main issues applying clustering techniques to a loyalty database are the sparsity in the data that turns the algorithm time-inefficient to process, and the uncertainty inherent to the assertive number of clusters selected to create the native baskets.

Price Index: Methodology Overview

Price index is interpreted as how good the prices in a retailer are compared to its competitors for a set of products (Galkin, 2019).

Generalizing from a stock-keeping unit (SKU) level to a broader level of granularity based on effective transactions tend to generate overestimated good price indexes. This limitation stands out when price changes and we experience differences in price indexes directly affecting the sales made by highly price

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

sensitive customers, with lower representation of products with prices above competitors and higher representation of products with prices below competitors in a specific period.

To avoid getting the same weight between competitors that might have significantly different impact to customers, one approach is to do it based on national market share. The limitation is that it considers the same importance of all competitors to different customers, despite if shopping at those competitors is possible or not (most Portuguese retailers do not have online stores or do not cover deliveries to all regions).

RETAIL-BASED FINE-TUNING APPROACH

Native Baskets: Product Granularity

The native baskets are intended to reflect a customer's purchasing list as accurately as possible. An important part of this step is the granularity a customer considers when thinking about the items on the list.

Intuitively one knows a customer does not think at a stock-keeping unit level but how to define the appropriate level of listing product? E.g. does a customer buying Basmati Rice Own Brand 1 kg takes into consideration the rice attribute alone or is the type of rice prevailing when planning a purchase?

Many studies focus on creating conceptual hierarchies modeled per instance using directed acyclic graphs (Gouider & Farhat, 2010), but the sparsity of the dataset on the SKU level is hardly mentioned. Rolling up rare items to higher levels in a product hierarchy prevents them from being excluded from the analysis and virtually makes them more frequent.

In addition to helping generalize items, product hierarchies are commonly created under a structure aiming to support day-to-day business operations.

Whenever a product hierarchy, built with taking the customers view into consideration, is already in place in the company its use is recommended over a traditional product hierarchy having suppliers in mind. Whatever the product hierarchy, the challenge regarding the product level remains the same since those levels are not meant for taking into consideration the customer substitution.

The authors will refer to this hierarchy on a customer view as virtual product hierarchy (VPH), fully independent from the commercial product hierarchy. This VPH allows, e.g. to create a virtual product of orange juice combining freshly squeezed orange juice and industrial packs of orange juice. The general structure of VPH is shown in Figure 1, from VPH level 1 as a category with a broader level to VPH level 4 as product-brand which is the more granular.

In order to fine-tune the optimal level on different VPH there are two main stages:

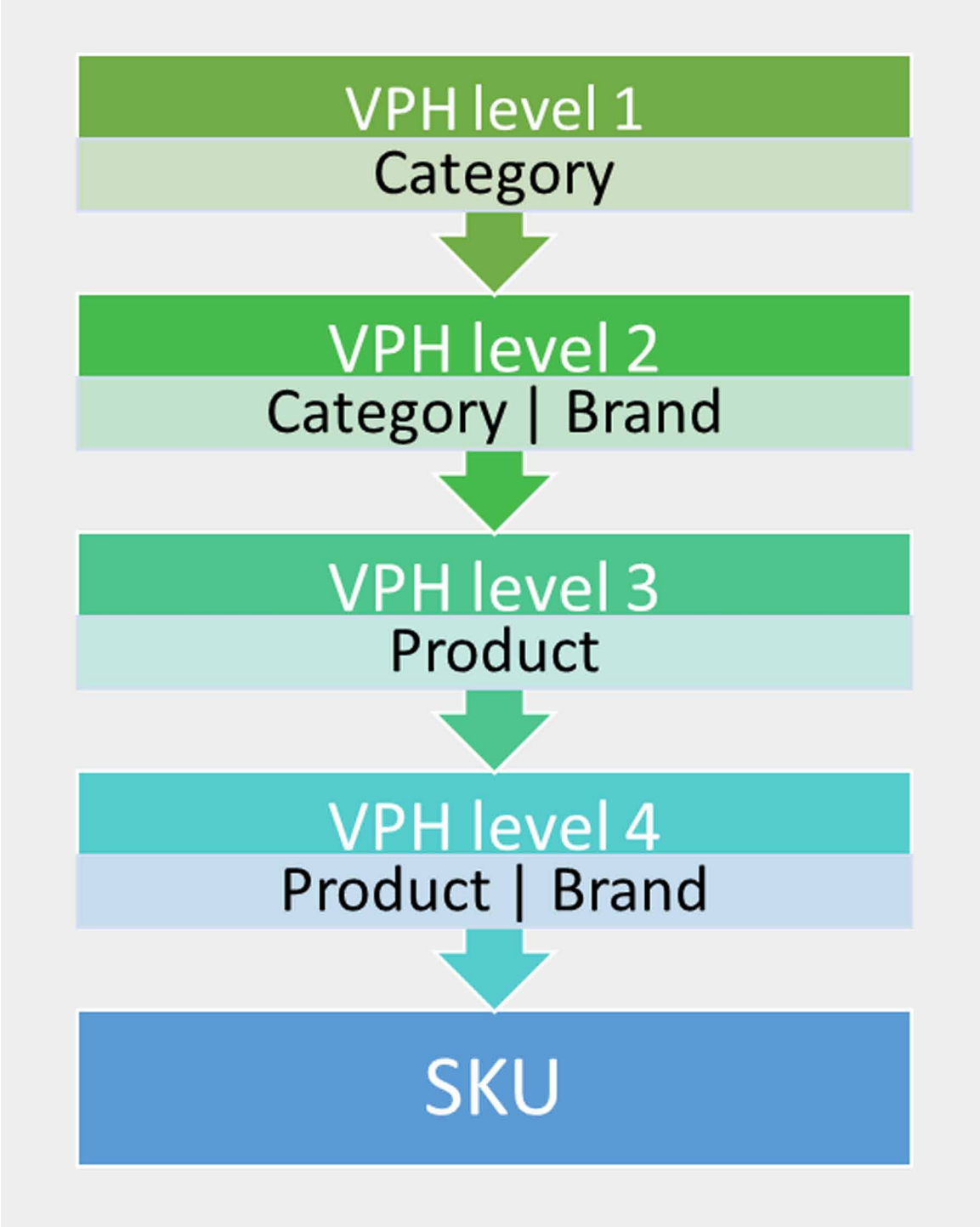
Step 1: Clean virtual products to consider
Step 2: Explore the substitution effects between products from the same level

In Step 1, business strategy defines that only daily products (business areas with high purchasing frequency) are to be candidates for the native baskets to track the price index.

Food products, hygiene, beauty & cosmetics, household care (e.g. detergents, toilet paper), pet food & care will be considered. Whereas textiles (fashion for both kids and adults, home textiles such as bedding and tablecloths), homeware (e.g. kitchen utensils), stationery (e.g. office and school supplies), culture (e.g. books) and electronics (e.g. kitchen appliances, heater) will not.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. VPH structure



Highly seasonal products comprised in business areas with high frequency were intentionally not excluded since we want to be able to detect seasonal native baskets, given the importance of price per-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ception on these sorts of products.

On Step 2, to fine-tune the product granularity to be considered one is interested in keeping a level of products that show low substitution between products within-category (VPH level 1). In terms of interpretation, products should be grouped whenever they share common attributes from a customer perspective. Levels of substitution may vary from category to category.

Let substitution be defined by the equation in the formula below:

$$Substitution\ Effect\left(S\right)=\frac{\#\,customers\ buying\ \ products\ i\ and\ \ j\ always\ in\ different\ baskets}{\#\,customers\ buying\ at\ least\ one\ of\ \ the\ two\ \ products\ i\ and\ \ j}$$

In Figure 2, one can see an example of substitution between products, where baby wipes from two different brands show higher substitution than baby wipes and makeup remover wipes.

Figure 2. Substitution business interpretation example of baby wipes

Two product groups are considered substitutes when they are recurrently bought by the same customer in different tickets. Product groups bought in the same ticket are interpreted as complementary products.

The substitution was computed between VPH of level 4 within the same VPH level 1.

The computation of the substitution on a higher level than 4 (levels 1, 2, and 3) is the result of a weighted average of the substitution degrees from the immediate lower VPH level.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

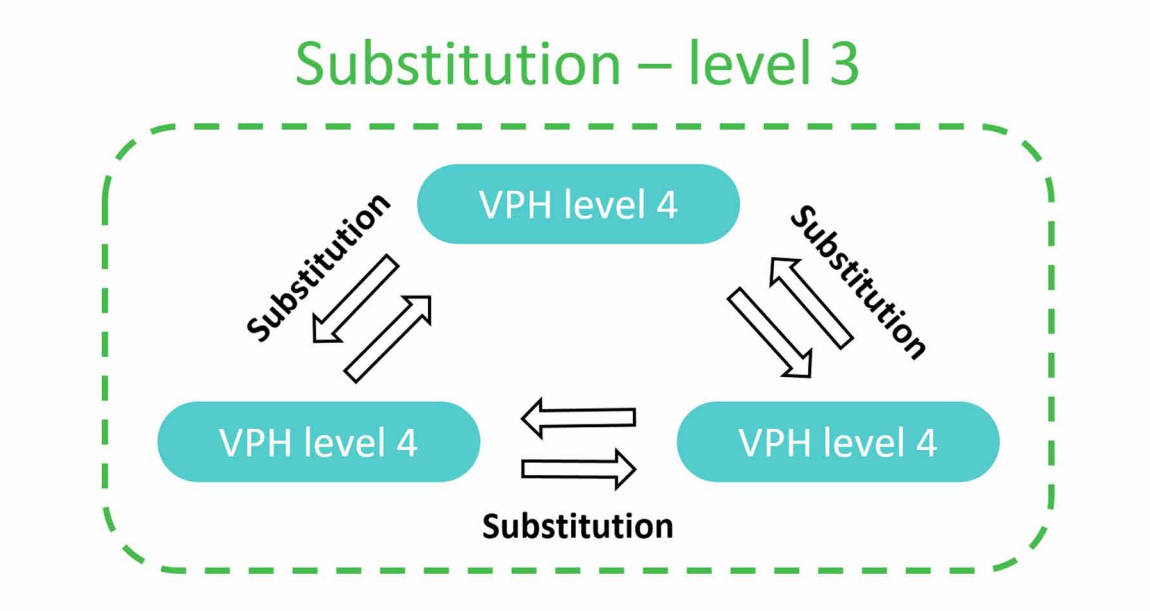
In Figure 3, one shows the computation scheme for the substitution degree of VPH level 3, as an example. The substitution of a given product group on a VPH level 3 is the average of the substitution degrees from the correspondent VPH level 4 product groups weighted by their number of customers, as described on the generalized formula below:

$$Substitution\ Effect\left(S\right) Prod_{level} = \frac{\sum_{l=1}^{n}\left(S_l \times \#customers_l\right)}{\sum_{l=1}^{n} \#customers_l}$$, n as all combinations within Prod-

level

The weighting average will balance the final substitution degree in cases with a high substitution degree between two product groups but very low representation in terms of customers who exhibit this behavior.

Figure 3. Computation scheme for the VPH level 3 substitution degree

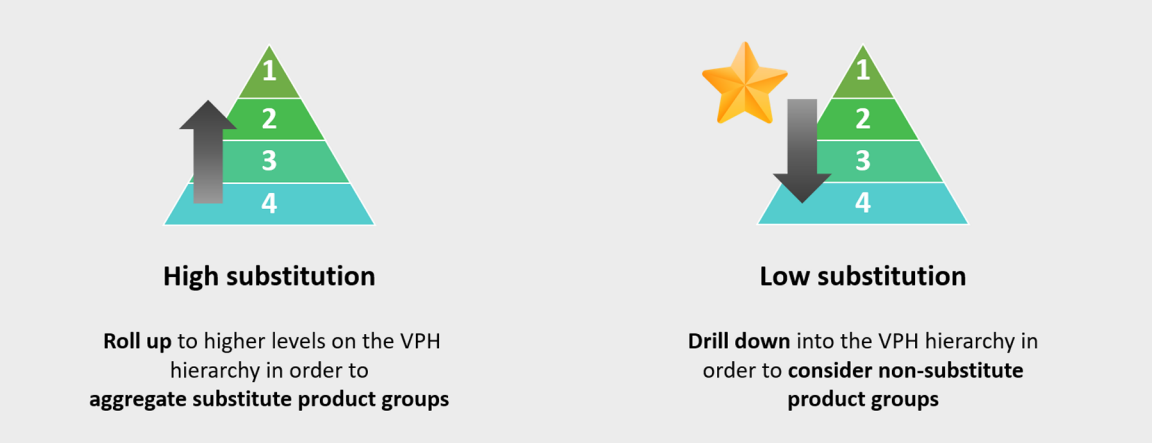


On a practical note, in the presence of product groups with high substitution one consecutively rolls up to higher levels of the VPH hierarchy. In the presence of low substitution between product groups, one may need to drill down into a lower VPH hierarchy in order to consider non-substitute product groups on the native baskets as shown in Figure 4.

In a business interpretation, products with high switching with each other shouldn't both be present on the same list/native basket, and thus should be grouped by rolling up to the immediately upper level of VPH. On the other side, one should ascertain that a group of products with low switching is not composed by two or more product groups not switchable with each other, and thus should be split by drilling down into the immediately lower level of VPH.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Substitution action scheme



The authors recommend starting from the higher level on the VPH (category as VPH level 1) and drill down into lower levels of the VPH whenever there are evidences from the substitution degree.

A tool has been created with the main metrics to enable the decision on the optimal level on the VPH to define the native baskets (Figure 5). For each VPH level 1 – category – the tool has two different types of metrics:

- Business metrics that allow the business user to understand how important a VPH is to the retail store through: percentage of tickets that the VPH is present, percentage of customers that buy the products within de VPH group, number of SKUs that the VPH group is made of (to understand how much granularity is within the VPH group) and the percentage of sales the VPH represents.
- Substitution metrics, such as the weighted substitution effect, as described above, the substitution effect distribution of all VPH within. Lastly, to guide the decision to drill down, or not, the VPH level, as described in Figure 4, a substitution index is presented:

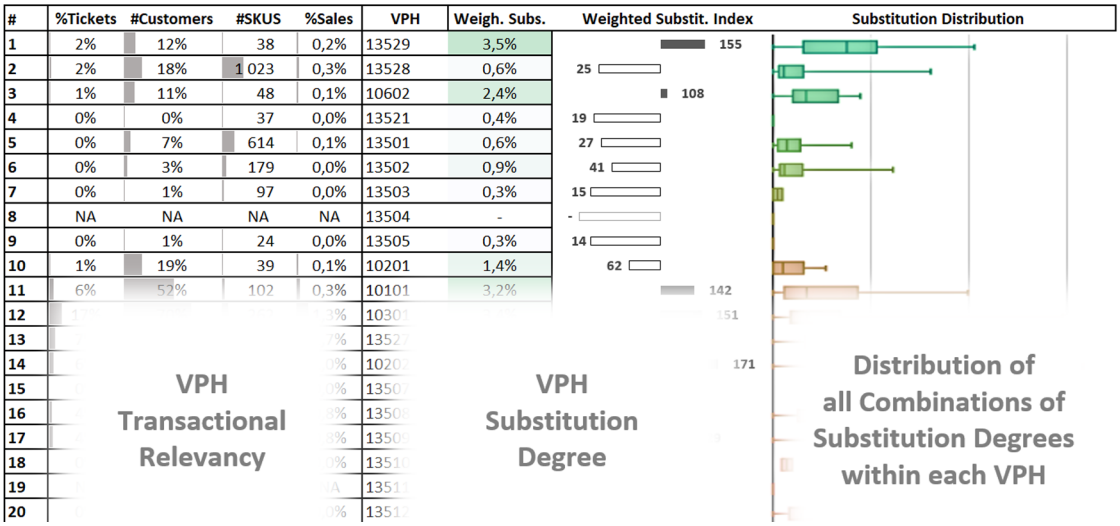
$$Substitutions\ Index = \frac{\text{Substitution Effect of the VPH group}}{\text{Average Substitution Effect of all combinations}} \times 100$$

A substitution Index below 100 indicates a below average substitution effect, meaning that VPH level 1 has the potential to be lowered. One should also consider a minimum relevancy of sales, transactions and customers to ensure meaningful listing products to lower the VPH level.

In the end, the tool should help one to define the most adequate level for each group of products to be considered in the construction of the Native Baskets.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Product level for VPH fine tuning tool



Native Baskets: Combining Customer DNA and Purchasing Moment

The native basket can be described as the intended shopping list of a group of customers in a given scenario. One must keep in mind that a customer might have different shopping lists/native baskets depending on their shopping mission and their buying profile.

The main assumptions regarding native baskets are their uniqueness among each other and coverage of the overall tickets in store and the behavior of customers. An experiment was conducted to find the pair of type of customer and type of transaction that yields the most diverse typology of tickets, in which the native baskets would be computed – a certain type of customers with a certain type of transactions would result in the creation of a set of native baskets.

Regarding the type of customers, the authors used three customer segmentations already implemented in previous works:

- Lifestyle segmentation – classifies a customer as being healthy-minded, family-driven, or price bound, based on the type of products bought
- Life stage segmentation – classifies a customer as being active adults, family with kids, or as seniors, based on the type of products and quantities/volumes bought of each product
- Price sensitivity segmentation – classifies a customer, based on how sensitive it is to changes in promotions/price and brand switching, as focused on every-day low prices, focused on a good deal/promotion or as price-upscale

Regarding the type of transactions, the authors used:

- Information of shopping mission – classification as stock-up transactions (monthly transactions – a large and comprehensive grocery basket), fill in transactions (weekly – a small basket with a

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

focus on fresh products or with promotion in mind), or food for now (occasional – small basket with focus on products for immediate consumption – bread, deli, meals, and drinks).
- Information of the week – weekdays versus weekend
- Information of the time of day – morning, lunch, early afternoon, late afternoon, night
- Information on the type of store format – hypermarket store, supermarket store, proximity store

The Euclidean distance between two combinations of type of customer and type of transaction A and B is given by the formula below, where a and b are the percentage of transactions VPH i is in transactions from all combinations A and B. Let n represent the number of all VPHs considered for the distance.

$$Euclidean\,distance(A,B) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

For each type of customers, the Euclidean distance was then weighted by the number of transactions in each type of transactions. The result can be found in Figure 6, where the bigger the Euclidean distance the more diverse are the transactions.

Figure 6. Weighted average of Euclidean distance between types of customers and transactions

Transaction

Customer	Shopping mission	Day of the Week	Time of day	Type of store
Life Style	0,61	0,30	0,24	0,29
Life Stage	0,73	0,24	0,17	0,22
Price Sensitivity	0,52	0,22	0,19	0,22

Several conclusions can be drawn from the experiment above. The authors concluded that the time of day is the least impacting element on heterogeneity and shopping mission the greatest impacting factor, followed by the store format. Meaning that transactions are more likely to be different depending on their shopping mission or day of the week than the time of day.

In terms of customer DNA, the lifestyle of a customer is the leading factor in the heterogeneity of transactions on the different types of stores. This result is a reason to believe that the choice of store format is most impacted by the lifestyle of the customer and not its price sensitivity or stage of life.

The life stage in combination with the shopping mission results in the most diverse set of transactions. This is the ideal scenario to create native baskets to guarantee full coverage of behavior in the retailer.

In conclusion, as stated above, the native baskets will be computed on the most diverse set of combinations of customer and transaction information: life stage segmentation with shopping mission.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Native Baskets: Create Native Baskets

As stated by the authors on the previous section, more specifically on *Native Baskets: Methodology Overview*, there are three methodologies to achieve native baskets – the list of shopping intentions of customers: through predictive algorithms, market basket analysis, and clustering analysis. In this section, the authors explore the viability and implementation of these methodologies to create native baskets.

Predictive Algorithms

The use of predictive algorithms introduces a personalization aspect to the creation of shopping lists – generating a prediction for each customer. However, the authors point out two limitations for this approach to the construction of native baskets.

Primarily, a predictive algorithm such as TBS, is based on the premise that a set of input data of a customer gives one list of products as an output. The purpose of this research is to create diverse lists emulating different buying patterns of the same customer/group of customers, making this approach not appropriate.

Secondly, the works mentioned above are biased towards the actual behavior of the customer which is influenced by promotional activity – if a set of products is heavily discounted in other retail stores, the customer might change his/her shopping list or even switch retailers – making it unviable to create generalized native baskets.

Market Basket Analysis

Another method for the creation of native baskets is based on a Market Basket Analysis methodology of pattern recognition, Frequent Item Set (FIS) mining.

FIS algorithms tend to be heavy on memory usage when dealing with large transactional databases. Taking into consideration this performance challenge, it is recommended to filter the frequent item sets through maximal closed item sets, resulting in fewer overlap buying behaviors. As an example, when one gets a set of 3 items as frequent and another set comprising the same 3 items plus a different one which is also frequent, one only retains the last maximal FIS.

Two algorithms of FIS were tested, using the software R library "arules", apriori and eclat. The latter proved to be more time-efficient, making it the prime candidate to generate FIS in transactional data.

For this implementation, the authors defined both support – the minimum number of occurrences of the FIS – and the minimum length of the set.

The minimum support was considered as 0.5% of the number of customers in the transactional database – meaning that a FIS would represent the behavior of, at least, half percent of all customers considered. As the minimum length of the FIS, the authors considered the first quartile of the distribution of the number of products per type of customer and type of shopping mission.

The FIS mining was applied to one of the combinations of life stage segment – families without children – and shopping mission – stock up mission with a high variety of products. Nevertheless, the maximum length of the FIS was 7 items, which is a shy number compared to what a typical stock-up mission entails.

In order to use the pattern mining results, the authors tested a clustering analysis on the FIS. However, with a thorough interpretation of the clustering groups obtained, the results were not relevant and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

did not convey the native basket nature that the authors intended – the clusters resulted in very similar groups of itemsets within each other.

Clustering Analysis

In this section, the authors will discuss the several stages that encompass the creation of native baskets using clustering analysis in transactional data. The purpose of this approach was to find the different clusters of transactions and extract the core behavior of each cluster – the core products or native baskets.

The first stage of the approach was to prepare and cluster the transactional data of customers.

To work with transactional data on clustering algorithms, linear data tables were transformed into a sparse matrix where each row corresponded to one transaction, filled by binary flags of 0 – non-bought items – and 1 – bought items – for each VPH category defined in the current Section, more specifically in *Native Baskets: Product Granularity* (Table 1).

Table 1. Binary representation of the database transformation prior to perform clustering analysis

	VPH A	**VPH B**	**VPH C**	**VPH D**	**…**	**VPH m**
Transaction 1	1	0	1	0	…	0
Transaction 2	0	0	0	1	…	0
…	…	…	…	…	…	
Transaction n	1	0	0	1		0

One of the main problems using sparse binary matrixes' transactions in a clustering algorithm is the fact that a large number of non-bought items could result in a closer relationship between transactions than bought items. To overcome this issue, the authors use a distance metric that would not consider matching zeros as a closer distance between transactions. For this reason, the Jaccard Index was applied as the distance metric to build the clustering analysis.

$$Jaccard\ Index\left(A,B\right)=\frac{\left|A\cap B\right|}{\left|A\cup B\right|}$$

The Jaccard index is used to measure how much of an overlap there is between two sets A and B. The measure to compute similarity between transactions should not be influenced by price, quantity or volume bought of each product, but simply by the presence or not of a product in both sets, hence the binary representation of bought (1) or non-bought (0) items.

Considering the wide range of clustering methods available, the authors opted for hierarchical clustering as it proved to be more flexible to adjust the number of clusters intended. The number of clusters was optimized using the pseudo t squared statistic value – choosing the number of clusters as the value which yields a noticeable drop in the pseudo.

The process of fine-tuning the number of clusters was replicated for all the combinations of life stage segments and shopping missions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The second stage of the approach was to extract the core elements of each transaction.

The rationale behind native baskets is to find the intention behind the effective purchase of the customer. The authors argue that the core products of a group of transactions will be the most representative in that group. Rephrasing the same idea, products with a low presence on the clusters could be interpreted as superfluous or impulse purchases and not the main buying intention of the customer.

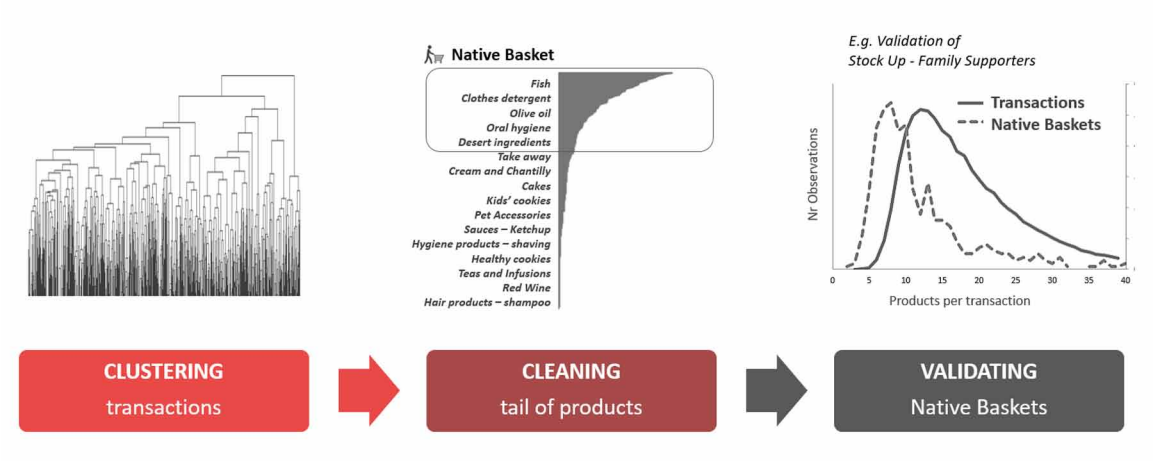
Taking this idea into account, for each cluster the percentage of transaction in which each VPHs was present was computed – hit rate. VPHs with a high hit rate are considered the core products of the cluster. Defining a threshold of hit rate, one can determine the core products of each cluster for each segment and shopping mission – the native basket.

The minimum hit rate to clean the tail of products from the native baskets was set to 40%. If a product was present in more than 40% of the transactions of the cluster, it was considered as a listing product to be part of the native basket. This fixed threshold was determined by optimizing the third stage of the method: the validation stage.

On the validation stage, the authors optimized the hit rate threshold to create native baskets with a similar length to the transactional data. However, the authors point out that a native basket, by nature, will usually be shorter in length than the actual transaction of the customer, as it only focuses on the buying intention and not on impulsive buying patterns.

As seen in Figure 7, although the distributions of transactions and native baskets are similar, in shape, their skewness differs. The native basket distribution skew will usually be greater than the effective transactions one. As native baskets are comprised of core products, their length is expected to be smaller than the transactions they represent.

Figure 7. Computation process of creating native baskets for each combination of life stage segment and shopping mission



In order to assign each transaction to a native basket, one can calculate the distance between the transaction and all native baskets from the same type of customer and shopping mission. The native basket with lowest Jaccard distance to the transaction is considered to represent the intent of the shopping visit.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Allocate Price Index

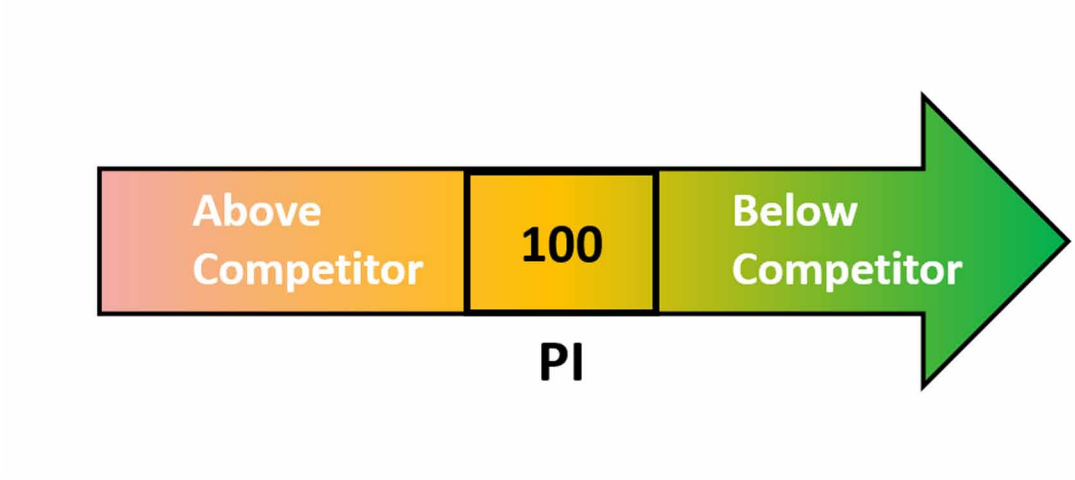
The promotional activity in Portuguese retailers is highly pressured.

In 2018, according to a market research from Nielsen (2019) 46% of all sales of domestic retail goods bought in Portugal had discounts.

With generalized weekly promotions allover retailers, it is fundamental that prices are picked on the competitors on a weekly basis, generating weekly price indexes per product (Figure 8) according to the formula:

$$Price\ Index = \frac{Retailer\ Price\ of\ a\ Product}{Competitor's\ Price\ of\ a\ Product}$$

Figure 8. PI concept per SKU and competitor



Prices on competitors are frequently picked by companies on the most important items in terms of contribution to total sales, combining both frequency and price per unit. Innovative products may be under price awareness, being flagged as products whose price should be monitored.

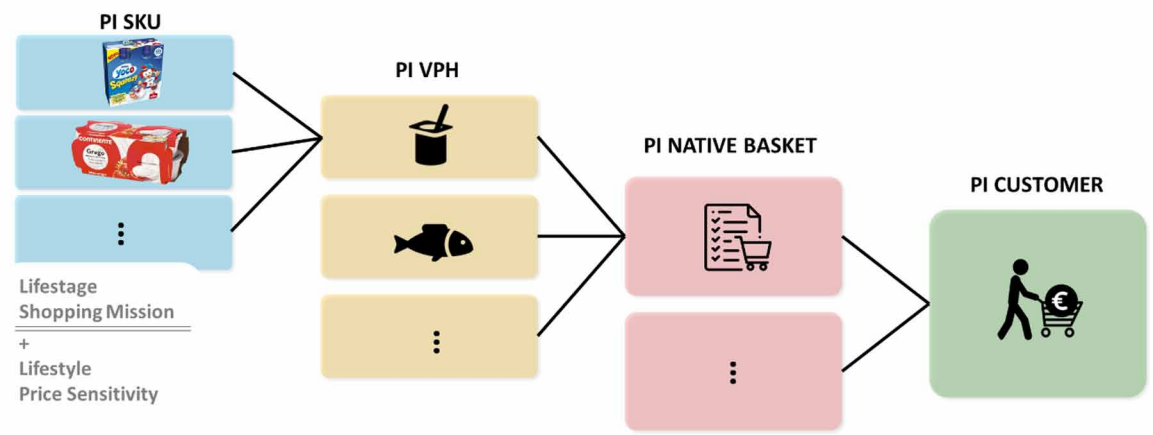
The price index aims to reflect the value proposition of a competitor against all others on the recurrent baskets of a customer – native baskets. In order to do so, the authors propose one should consider the different importance that distinct customers give to certain SKUs, products, and even native baskets.

Weighting the price index from the SKU to the desired level of aggregation becomes more of a business decision than a state of the art one.

It is crucial to involve the business on the desired metrics (number of customers, number of tickets, sales, quantities, volumes) to weight the price index from SKU to VPH to Listing Product to Native Basket to customer, always taking into consideration the type of customer under analysis. The process involves four steps per competitor as described in Figure 9, and comprises specific time periods, customer DNA families, and metrics. An extra step 5 aggregates the price index per competitor on an overall customer price index against all competitors.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Framework of the propagation from SKU price index to customer price index per competitor



Step 1: SKU PI

The SKU PI is collected on a weekly basis, per competitor, and does not vary between customers. It feeds a database which periodic refresh, coverage and accuracy are fundamental into having an accurate customer price index estimation.

Step 2: VPH PI

The importance given to a SKU within a listing product (VPH of desirable level for the native basket list) may be very distinct within customers from the same life stage segment and on the same shopping mission. The price sensitivity (how important price variations are in changes to the purchased quantities) and the lifestyle (importance given to healthy, pre-packed and pre-seasoned, economic products) of a customer are factors that affect the group of SKUs a specific type of customer considers within a purchasable product.

At this specific step, such segmentations are taken into consideration to weight the SKUs PI purchased in the previous three months (retailer stability period) within each listing product/VPH, by 50% sales and 50% number of customers.

Extending the period to capture the representation of each SKU on the listing products PI from one week to the stability period intends to mitigate the impact that possible weekly promotions or high prices might have in the overall importance of each SKU in sales, as while it allows seasonal patterns to stand out on SKUs.

Step 3: Native Basket PI

The representation of each listing product/VPH PI on the native basket will be affected by customer life stage and shopping mission weighting them by 50% sales and 50% number of customers buying each VPH in the previous three months.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Step 4: Customer PI per Competitor

The percentage of customers buying each native basket is the only metric affecting the weighted customer PI per competitor, taking into consideration the previous three months per life stage and shopping mission segments.

Step 5: Personalized Customer PI

One of the main steps is yet to be taken into consideration. At this stage, the PI per customer per competitor was successfully reached, yet not all customers will be impacted by all competitors with the same strength.

Thus, the authors propose to adjust the Huff gravity model methodology to balance the importance of each competitor PI to each individual customer.

The authors propose:

- main competitors in the customer trade area will be taken into account to compute the overall price index
- the store size of a competitor has influence on customer price index
- the distance does not impact customer perception with a linear distribution, but with a bell shape centered on the customer
- whenever more than one store from the same competitor is in the trade area of a customer the minimum distance will be measured.

A laboratory experience allows to establish the balance between store size and distance to stores on its impact in sales.

The formula proposed by the authors to compute the customer overall PI gives more importance to stores nearer to the customer and to competitors with higher area of sales (store size) as follows in Figure 10, considering only the store from each competitor nearer to the customer and within its trade area.

$$weight\ PI\ competitor_i = \frac{\frac{\sqrt{store\ area_i}}{\left(distance_i \times 100 + 1\right)^{3/2}}}{\sum_{i=1}^{n} \frac{\sqrt{store\ area_i}}{\left(distance_i \times 100 + 1\right)^{3/2}}}, i\ competitors$$

The final customer PI is completely customized on a one-to-one approach.

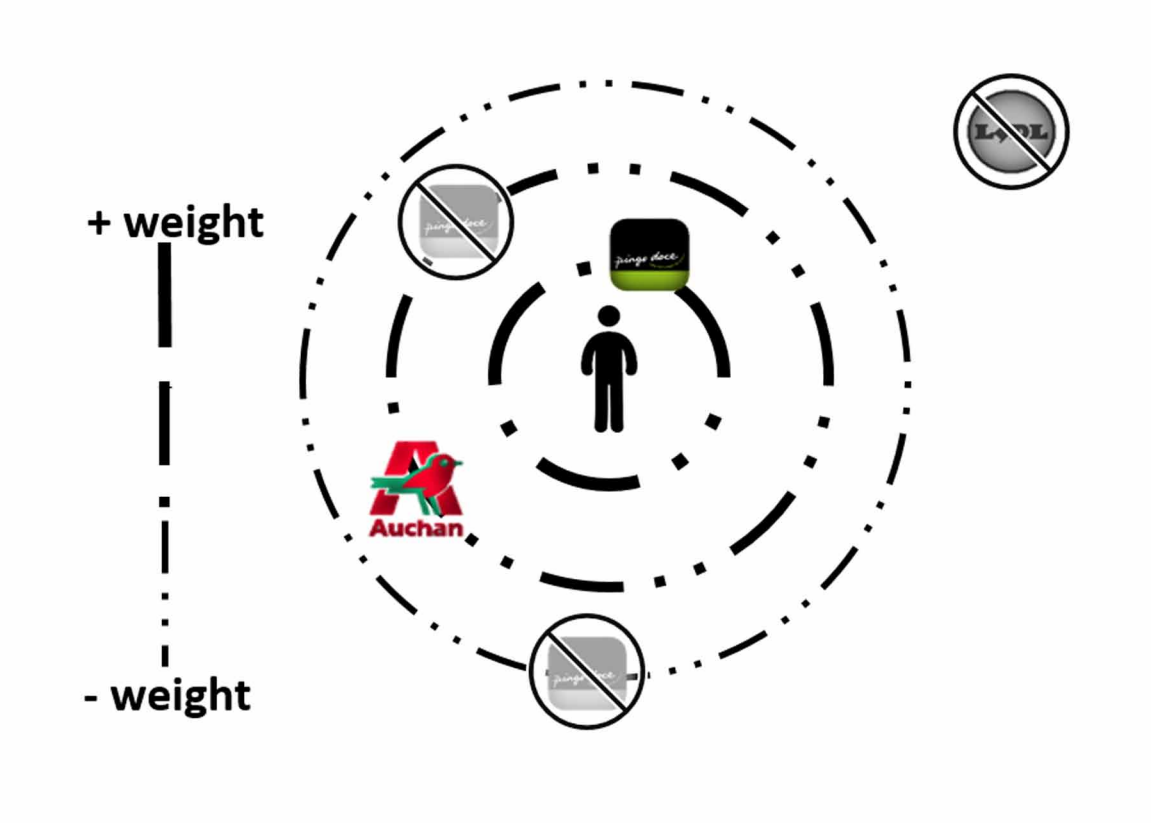
Stopping at step 4 is preferable whenever one wants to react specifically to a competitor and track the evolution of life stage segments, usually reported by market research consultants.

Implemented System and Business Application

The main steps of the methodology comprise building the native baskets (1) and defining the methodology to allocate the prices collected to a customer (2).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Weight of customer PI per competitor to compute personalized customer PI



The native baskets stage (1) covers:

- defining the level of product hierarchy to consider when the customer fills the native baskets/shopping lists – using a substitution metric
- fine-tuning the right combination of customer DNA / shopping characteristics that lead into more different native baskets among each other but less dissimilar within each group
- aggregate similar tickets categorizing each ticket as a native basket – using clustering techniques
- define the items on the native baskets to be under price index measurement, combining sufficient spread versus the total spent in store and the minimum products in order to discard impulse/less represented items on the basket

Having categorized all customers' native baskets into its effective tickets, allocating the prices collection stage (2) covers:

- considering the prices collected in that specific period (week) comparing to the main competitors
- moving from the item written in a shopping list to the specific product bought one must consider a lower granularity that takes into consideration customer preferences (e.g. include product price range, brand loyalty, promotion sensitivity, household packaging needs)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- allocate the weight of the different native baskets taking into consideration the stability period and natural frequency of customers and not only the tickets purchased in that specific week
- measuring the weight of the different items in each native basket can potentially be made using the number of customers, sales, frequency or volumes
- consider the distance/store area of each competitor on the trade area per customer and combine it to allocate a single price index per customer.

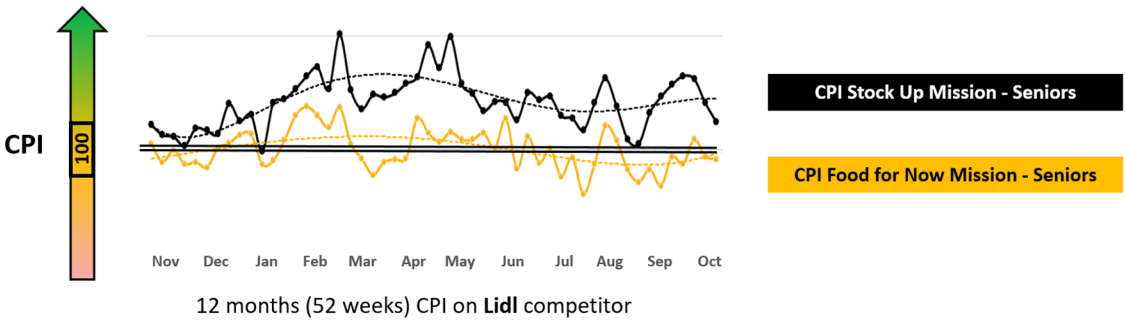
By not focusing only on the products bought per customer but also on similar shopping experiences from similar customers one can capture items that the customer is currently buying on competitors to measure its own price index.

This methodology allows to measure price index on a customer level enabling the personalization combining it with price sensitivity segmentations (e.g. create price awareness on non-sensitive customers instead of sending promotions), share of wallet segmentations (e.g. target baskets with poorly priced indexes in low share of wallet customers).

The authors create two distinctive tools to tackle two completely different approaches of turning the insights into action:

- Monitoring native baskets with consecutive poorly priced indexes allows the identification of combinations of products that are not balancing each other on specific segments of customers (Figure 11)

Figure 11. Tool to track customer PI on 2 focus shopping missions and 1 customer segment against one competitor



- Monitoring customer segments with consecutive poorly priced indexes allows tackling possible customer dissatisfaction on a one-to-one level that could potentially lead to churn.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

Customizing price per customer is crucial to prevent mass marketing promotions. The digital era is the big enabler for personalizing price, whereas it is not possible on physical purchases with physical display of price tags.

Using loyalty information, it is possible to combine customer DNA segmentations to target highly price sensitive customers in baskets with low price index, generating saves in costs.

This chapter presents room for improvements on the level on which the substitution degree was computed (always VPH4 grouped through weighted averages to roll up to higher levels of VPH). Whenever there are no time-consumption restrictions, the substitution effect could be computed by each individual level.

In terms of operational limitations, the methodology is highly sensitive to having good coverage on prices picked on the competitors of all items measured on the native baskets.

The proposed methodology is presented to measure Customer Price Index. The authors consider that the Native Basket Creation step could also be used to create typical shopping lists and could be applied to physical or online retail markets.

One of the main constraints of the framework presented in this chapter is the heavy reliance on loyalty data needed to ascertain the Native Baskets of customers. One possible future research could be to apply the approach to a transactional problem without a loyalty program and measure its effectiveness on capturing the intention of purchase.

Due to company policies regarding sharing sensitive information, the level of detail on the results presented by the authors is limited.

CONCLUSION

Personalized price index enables to customize communication to consumers depending not only on their relation to price but also on how competitive price is comparing to other retailers on the most relevant products to the customer. Competitor forces are also considered when personalizing price indexes, filtering nonexistent competitors on the surroundings of the customer and giving more strength to prices from competitors with higher proximity to the customer.

There are two main approaches that can follow from customizing price:

- mass approach: balance SKU price changes taking into account the impact on native baskets from specific segments or types of transactions (aligned with the business core strategies), or by the identification of native baskets recurrently not competitive in terms of price
- personalized approach: create strategies per combinations of DNA customer segments in terms of how to communicate the business position with regard to price directly related to the price index on the customers' core interests – native baskets –, or target customers recurrently under non-competitive price indexes in order to prevent churn.

As a long-term strategy, under a reality where physical store sales remain relevant and technology evolves enabling customers to compare prices at the distance of a click, pulling price index customization from online to mass or personalized approaches is a must to save costs.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors and was supported by SONAE MC, SGPS, S.A.

REFERENCES

Bouguessa, M. (2011). A Practical Approach for Clustering Transaction Data. In P. Perner (Ed.), *Machine Learning and Data Mining in Pattern Recognition* (Vol. 6871, pp. 265–279). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-23199-5_20

Doan, T., Veira, N., Keng, B., & Ray, S. (2018). *Generating Realistic Sequences of Customer-Level Transactions for Retail Datasets*. *IEEE International Conference on Data Mining Workshops*, Singapore. 10.1109/ICDMW.2018.00122

Galkin, A. (2019). *How to Calculate Price Index. Formula to Know Competitors' Impact on Your Sales in 15 Minutes*. Competera. Retrieved from https://competera.net/resources/articles/price-index-calculation

Gouider, M., & Farhat, A. (2010). Mining Multi-level Frequent Itemsets under Constraints. *International Journal of Database Theory and Application*, *3*(4), 15–34.

Guidotti, R., Monreale, A., Nanni, M., Giannotti, F., & Pedreschi, D. (2017). Clustering Individual Transactional Data for Masses of Users. In *Proceedings of the 23rd International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery. 10.1145/3097983.3098034

Guidotti, R., Rossetti, G., Pappalardo, L., Giannotti, F., & Pedreschi, D. (2019). Personalized Market Basket Prediction with Temporal Annotated Recurring Sequences. *IEEE Transactions on Knowledge and Data Engineering*, *31*(11), 2151–2163. doi:10.1109/TKDE.2018.2872587

Hustić, I., & Gregurec, I. (2015). *The influence of price on customer's purchase decision*. Academic Press.

Jean, W., & Yazdanifard, R. (2015). The Review of How Sales Promotion Change the Consumer's Perception and Their Purchasing Behavior of a Product. *Global Journal of Management and Business Research*, *15*, 33–37.

Kumar, V., & Leone, R. P. (1988). Measuring the Effect of Retail Store Promotions on Brand and Store Substitution. *JMR, Journal of Marketing Research*, *25*(2), 178–185. doi:10.1177/002224378802500206

Mahanti, A., & Alhajj, R. (2005). Visual Interface for Online Watching of Frequent Itemset Generation in Apriori and Eclat. In *Proceedings of the Fourth International Conference on Machine Learning and Applications*. IEEE Computer Society. 10.1109/ICMLA.2005.68

Nielsen. (2019). *3,5 Mil Milhões de Euros na "Selva Promocional" dos Bens de Grande Consumo*. Nielsen. Retrieved from https://www.nielsen.com/pt/pt/insights/article/2019/3-point-5-billion-euros-in-promotional-jungle-of-consumer-goods/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Prassas, G., Pramatari, K. C., Papaemmanouil, O., & Doukidis, G. J. (2001). A recommender system for online shopping based on past customer behavior. *Proceedings of the 14th Bled Electronic Commerce Conference e-Everything: e-Commerce, e-Government, e-Household, e-Democracy*.

Wang, J., & Karypis, G. (2004). Summary: Efficiently Summarizing Transactions for Clustering. In *Proceedings of the Fourth IEEE International Conference on Data Mining*. IEEE Computer Society. 10.1109/ICDM.2004.10105

Wang, K., Xu, C., & Liu, B. (1999). Clustering Transactions Using Large Items. In *Proceedings of the Eighth International Conference on Information and Knowledge*. Association for Computing Machinery. 10.1145/319950.320054

Wang, P., Guo, J., Lan, Y., & Cheng, X. (2014). Modeling Retail Transaction Data for Personalized Shopping Recommendation. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*. Association for Computing Machinery. 10.1145/2661829.2662020

Yan, H., Chen, K., & Liu, L. (2006). Efficiently Clustering Transactional Data with Weighted Coverage Density. In *Proceedings of the 15th International Conference on Information and Knowledge Management*. Association for Computing Machinery. 10.1145/1183614.1183668

Yu, F., Liu, Q., Wu, S., Wang, L., & Tan, T. (2016). A Dynamic Recurrent Model for Next Basket Recommendation. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery. 10.1145/2911451.2914683

ADDITIONAL READING

Gooding, P. (2011). Consumer Prices Index and Retail Prices Index: The 2011 Basket of Goods and Services. *Economic & Labour Market Review*, *5*(4), 96–107. doi:10.1057/elmr.2011.44

Mahdi, M. A., Abdelrahman, S. E., Bahgat, R., & Ismail, I. A. (2010). F-tree: An Algorithm for Clustering Transactional Data Using Frequency Tree. *Journal of Engineering Sector of Engineering Colleges - Al-Azhar University*, *5*(8), 101-123.

KEY TERMS AND DEFINITIONS

Basket: A set of products bought on a transaction.

Category of Products: A set of products that fulfill the same customer need.

Listing Product: A group of stock-keeping units that are on a customer mind when preparing a future purchase of retail goods. It is a level of a product hierarchy grouping a set of products with low substitution effect.

Native Basket: List of products to purchase prior to effectively shopping. It is the intended set of items needed prior to factors in-store that influence the actual purchase such as the price index.

Price Index: A weighted average of price in relation to a competitor or a set of competitors for a given product or product group on a given time period. A product or set of products with a price index

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

over 100 is interpreted as having a competitive price on the own retailer, a price index lower than 100 is interpreted as being more competitive in terms of price on the competitors on a given period of time.

Shopping Mission: The main motivation when shopping. Main shopping missions involve from stock-up missions often known as the monthly shop to convenience purchases related to suppressing a specific need (e.g., diapers, wine).

Stability Period: Minimum period needed to capture a customer behavior on a business.

Stock-Keeping Unit: A scannable bar code and is interpreted as a single product with individual characteristics such as type of product, packaging, volume, brand and eventual specific components.

Substitution Effect: Measures the switching between a set of items. Substitute products are often purchased in different baskets by the same customer and are capable of accomplishing the same need.

Virtual Product: A group of SKUs with enough similar properties to turn them into the same product on a customer view. The virtual product belongs to a virtual product hierarchy fully independent from the commercial product hierarchy.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 9
Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence

M Dolores Méndez-Aparicio
https://orcid.org/0000-0003-1787-8031
Open University of Catalonia (UOC), Spain

Alicia Izquierdo-Yusta
https://orcid.org/0000-0001-5894-6577
University of Burgos, Spain

María Pilar Martínez-Ruiz
https://orcid.org/0000-0002-5890-5174
University of Albacete, Spain

Ana Isabel Jiménez-Zarco
https://orcid.org/0000-0002-8980-6814
Open University of Catalonia (UOC), Spain

ABSTRACT

Data science has established itself as a discipline of study in the new technological paradigm of the 21st century. Customer data and metadata are not only unique opportunities for companies that are interested in collecting, processing, and elaborating useful information for customer-centric business management. They are also starting to feed other disciplines, such as industry and computer engineering, which are developing new ways of organising information and intelligent response, where significant technological advances in facial, sensory, text, voice, and image recognition are constantly emerging, with response capabilities ever closer to human thinking. From this continuous interaction emerges a new asset for companies: knowledge. Thus, science applied to data is consolidating as an opportunity for companies in customer loyalty and retention by being able to include a relevant aspect in human relations such as warmth and empathy, basis of marketing 5.0, oriented once again towards human relations, from a virtual existence and total digitalisation.

DOI: 10.4018/978-1-7998-6985-6.ch009

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The **competitive pressure** produced by the **globalisation** of the economy has focused on the client as a strategic objective of companies (Kohli, Jaworski & Shabshab, 2019). From the perspective of an unprecedented consumer society, knowing their profile and offering **personalised** products that meet their needs is the great challenge in the field of marketing today. In this race for **differentiation**, providing **experiential consumption** is where brands have found their great opportunity to stand out from the competition (Triantafillidou & Siomkos, 2014).

This change in the economic scenario has been possible thanks to the democratization of the so-called **new information and communication technologies** (henceforth ICT) where connectivity and knowledge have transformed the way of understanding interpersonal communications and the way of relating to companies (Torrent-Sellens, 2020). This technological paradigm focused companies on attracting new online customers until the end of the 20th century given the penetration capacity and attractiveness of the digital channel (Ha, & Stoel, 2009). However, this strong competitiveness generated in the offer, has simultaneously generated the need to retain and to build customer loyalty that allows to maintain and to assure their portfolio (Mascarenhas, Kesavan & Bernacchi, 2006). This is how the concept of Customer Experience was born, where anticipating, surprising and enthusing the user is the objective (DEC, 2020).

According to Cisco's statistical sources (2020), more than 1000 terabytes (1 TB equivalent to 1000 gigabytes, GB) of information is generated monthly on the Internet. This reputable IT consulting firm estimates that by the year **2022** there will be as much data traffic on the internet as has been accumulated over the last 32 years, with **60% of the world's population connected to the internet** (computers, mobile phones, tablets, watches or smart TVs); global mobile networks will record **more than 12 billion mobile devices** and IOT (Internet of Things) type connections. This huge amount of data will exceed the Zettabyte milestone (1 ZB equals 1 million terabytes, Cisco, 2020). This huge amount of information has overwhelmed the companies' information systems and made the so-called digital transformation necessary (Sarenet, 2019). In this new omni-channel environment, companies must look for new resources to attract their potential consumers and customers (Sabio & DEC, 2020). Being the first choice for the consumer requires intelligent browsers that can efficiently process increasingly complex searches within distributed and, in most cases, defragmented information. Similarly, designing the information architecture required by the new informed citizen implies that the technology must guarantee information that is not only current, but also adapted to our needs, sufficient and precise (Méndez-Aparicio, Jiménez-Zarco, Izquierdo-Yusta & Blazquez-Resino, 2020). Characteristics that can only be satisfied from **Applied Data Science** (hereinafter, ADS).

Thus, we are in the new **Knowledge Economy** (Hendarman & Tjakraatmadja, 2012, Torrent-Sellens, 2016) and the new **Behavioural Economy** (DEC, 2019), where intelligence applied to data has become a service for companies, information a right for customers, and innovation a brand attribute. This reality has been shared by business and industry and not in vain, companies are immersed in a new technological transformation with the irruption of 5G (Martín Lineros, 2020; Torrent-Sellens, 2020).

This chapter is set in this socio-economic context and aims to explain how technology applied to data can help in the sustainability of the economy, creating loyal customers from the information collected in each mutual interaction as a source of data. This knowledge not only allows the offer to be personalised and empathy with the service to be improved, but also makes it possible to create new innovative products and services using virtual reality technology, mobility and consumer behaviour models. Through innova-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

tive applications, companies can better anticipate customers' actions, as well as leverage the automated behavior of AI to match responses based on customers' personal profile and consumption situation.

In the first section of this chapter dedicated to Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence, the concept of customer experience, its definition, background, and consequences on the business will be developed. The second section will describe what **Applied Data Science** is. Each section will describe the different perspectives of studies, types of data generated, the evolution in information processing and how technology becomes metatechnology making automated decision making, or Artificial Intelligence, possible. In the final section, we describe the business implications in detail, not only in terms of consumption, but also regarding the type of relationship that companies can develop with their clients. After all, the intrusiveness of these technologies has led to serious reflections on the importance of freedom and individuality. Thus, in this section will discuss these concepts and how the law can best legislate on them.

In the first section of this chapter dedicated to Studying Customer Experience and Retention Using Applied Data Science and Artificial Intelligence, the concept of customer experience, its definition, background, and business implications will be developed. In the second section, we will describe what Applied Data Science is. Each section will describe the different perspectives of studies, types of data generated, the evolution in the treatment of information and how technology becomes metatechnology making possible automated decision making, or Artificial Intelligence. Finally, the business implications will be described in detail, not only in the consumption, but also in the type of relationship of the company with its customers. The intrusiveness that technology has advocated has led to serious reflections on the concept of freedom and individuality. In this section, we will also discuss these concepts and how the law has had to legislate on them.

WHY THE CUSTOMER EXPERIENCE?

It is necessary to go from a consumer experience to an extraordinary one, if the company wants to be remembered (see Figure 1).

That said, a report from the Boston Consulting Group (BCG & DEC, 2018) indicates that the expectations of customers are increasingly demanding in their **empowerment** as consumers and has **penalized loyalty with their brands**. The decline in loyalty is so marked that the customer recommendation index has fallen back to **2014** values despite the effort made by brands to satisfy and surprise their customers, as shown in figure 2. Bearing in mind that the recommendation is directly related to an **increase in revenue** (itUser, 2020), Forrester noted (2020) that improving the customer experience by one point can generate more than a billion dollars in additional revenue (with reference to the automotive sector, at least). This decline clearly has a serious impact on companies' profit and loss accounts.

As can be seen in this figure 2, the comparison is based on the brand advocacy index (BAI). BAI is an accurate indicator that measures customer experience based on actual customer and non-customer reviews and recommendations about brands. Although Net Promoter Score (NPS, Reichheld, 2014) reflects customer loyalty and their recommendation intention, BAI becomes a strategic indicator as it measures the real impact of the brand in the market.

In view of this situation, numerous investigations have been carried out on the determinants of customer expectations that shed light on consumer behaviour in the new socio-economic context (Méndez-Aparicio, Jiménez-Zarco & Izquierdo-Yusta, 2017). However, despite the effort of the theoretical frame

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1.
Source: (Carù & Cova, 2013)

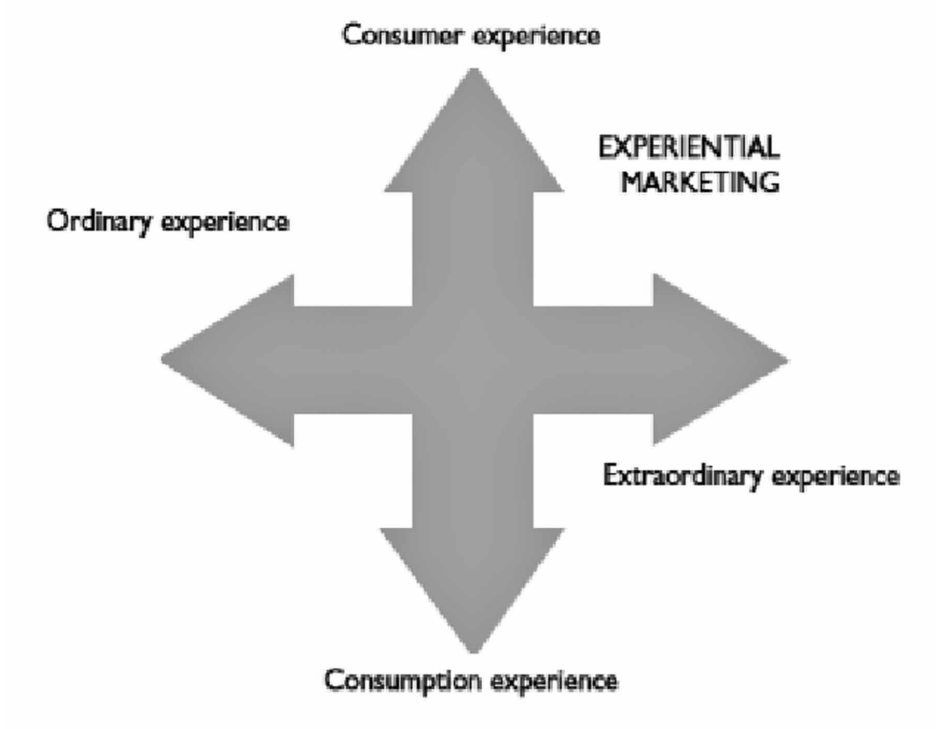


Figure 2.
Source: (Boston Consulting Group, BCG & DEC, 2018)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of reference, there are few studies that include complex behavioural variables or scientific methods beyond mere statistics, as indicated in their report by the authors Pérez-Rave and Muñoz-Giraldo (2014). Thus, of their comparison of publications in accredited journals (original in SciVerse Scopus) only 22% of the works consider the **attitudinal** aspect of the customer-company relationship and only 46% of the researches used **causal models** (SEM, Structural Equation Model) to confirm their hypotheses.

In the following sections of this section, we will describe in more detail what the customer experience is and how it is achieved, as well as the strategic importance of the customer experience in the retention and sustainable growth of the company.

What is the Customer Experience?

The most widespread definition of the customer experience is in terms **of memorable and surprising** (Oh, Fiore and Jeong, 2007, p.120). The Spanish-speaking community has usefully defined this concept as a **"wow moment**" (although not officially recognised by the *Real Academia de la Lengua Española*, RAE, 2020), which translates to **"surprise that pleases".** Therefore, although the experience can be positive or negative, the wow moment is understood as one that exceeds the customer's expectations and causes full satisfaction and a **memorable memory** (Tokman, Davis, & Lemon, 2007).

Similar definitions follow one another in the reference framework. Thus, Walls, Okumus, Wang and Kwun (2011) corroborate this definition by saying that the experience is the **impression** the customer gets, whether positive or negative, from consumption. These authors emphasise the term impression, referring to **impact,** as also argued by Pine and Gilmore (1999, p. 12) and that the consumer experience involves the user in a personal way, which makes his or her memory special. This impact on the customer is possible because the customer experience happens in an integral way in people, **both in its cognitive, affective, social and physical aspects** (Lusch and Vargo, 2014) and where the customer appreciates the **hedonic** aspect of the transaction but also the **quality and perceived utility** of the product and service he or she consumes.

As can be seen in table 1, the research framework has typified the different dimensions of the customer experience (see table 1) considering the response obtained from the sensory (degree of participation of the senses in the experience), affective (feelings outcrop), cognitive (knowledge generated), physical

Table 1. Dimensions of the Customer Experience

Authors	Dimensions (types of response given)
Schmitt (1999)	cognitive, sensory, affective and relational
Triantafillidou & Siomkos, (2014)	cognitive, emotional and behavioral
Zomerdijk & Voss (2011)	cognitive, sensory, emotional, pragmatic, lifestyle and relational
Verhoef, Lemon, Parasuraman, Roggeveen, Tsiros & Schlesinger (2009)	cognitive, affective, emotional, social and physical
Same & Larimo (2012)	cognitive, sensory, affective and behavioural
De Keyser & Lariviere (2014)	cognitive, sensory, emotional, physical, spiritual and social
Lemon & Verhoef (2016)	cognitive, sensory, emotional, behavioural and social
Adhikari & Bhattacharya (2016)	cognitive, sensory, emotional, physical and relational

Source: (Méndez Aparicio, 2019)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

(actions carried out) and relational (socialisation and identity established), emotional, pragmatic and lifestyle perspectives.

In short, the different manifestations of the customer's experience referenced in the research framework respond to the evolution of consumer **expectations** (Walls et al., 2011). Thus, from the initial intuitive concept that the experience is based on its hedonic capacity, there really underlies the reaction of emotional involvement that this experience has provoked in the user and that makes it remain in his **memory** (Pelegrín-Borondo, Arias-Oliva & Olarte-Pascual, 2017). This assertion corroborates, once again, its strict **personal character.**

This argument is especially valid in light of COVID-19, which has radically changed the customer's value range (Findasense, 2020). To respond to this challenging (albeit temporary) context, companies and researchers must investigate "how" to exceed expectations.

How the Customer Experience is Achieved: The Customer Journey and the Voice of the Customer

For an experience to be memorable for the customer it must be based on **quality** and must **provide value** (Swaid & Wigand, 2012; Chang, Kuo, Hsu & Cheng, 2014) in: 1) the quality product or service, which must be outstanding to generate its memory and 2) through a quality service channel, which makes it easier for customers to use the company's services.

But there is no doubt that a relationship is forged in everyday relationships and will depend directly on the strength and interest given to them (Spence, 2019). Maintaining a warm, continuous and cordial interaction, which consolidates trust and involvement, has also become a way of surprising the client. Consolidating this relationship throughout all the points of contact with the client (physical or intangible), or **Customer Journey**, is a new way of building customer loyalty, which thanks to the Enterprise Feedback Management Systems (hereinafter EFMS) can be managed.

This new emotional value of the experience is made possible by **Applied Data Science,** which allows for unprecedented personalization. Thus, interactions between the customer and the brand that have not traditionally been considered relevant have begun to show their value as they are significant for customers (Triantafillidou & Siomkos, 2014; 6th DEC Congress, 2019). For example, they now pay special attention to the way in which a product or service is delivered, the digital accompaniment, and that companies actively listen to their opinions and complaints (Balaguer, 2020). These interactions have become an excellent experiential communication opportunity. Moreover, these situations or **"micro moments",** as Google calls them (Direct Marketing, 2019), can be as relevant as the object of consumption (Méndez Aparicio, 2019). This continuous relationship with the client is increasingly less linear and more experiential, where communication becomes a true and continuous interaction (Alonso-Cortés, 2020).

As can be seen from the above arguments, anticipating and making a repeatable customer experience, with a vocation for self-improvement, is only possible from an excellent knowledge of the customer, of his or her groups of influence, as well as of the psychographic context of his or her consumption. Hence, the customer experience can be defined as the memory generated in the consumer as a consequence of his or her relationship with the brand (DEC, 2017). These interactions occur continuously in the relationship between the customer and his company, a situation that has increased with the exponential growth of **digital touchpoints,** where the physical and digital relationship coexist and are fed back into each other (Direct Marketing, 2019).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

This relationship can be direct with the brand or **point of contact**, or through the entire value chain that the brand generates in the transaction, the so-called **"moments"**. It is relevant to consider that the experience also involves intermediaries and retailers.

It cannot be said that there are different customer experiences, it is the different moments of the customer that make the difference. Finding these moments is only possible through research, either **quantitative** on the data generated by each contact or **qualitative**, through the **Voice of the Customer** (hereinafter **VoC,** or set of opinions that the company collects after each interaction), by inferring statistical models. Only in this way is the brand capable of knowing the importance that each customer attaches to each moment and how the particular experience they have felt has been, from the extreme values found to each of these questions. Thus, the "**moments of truth"** are those that generate the greatest memory in the consumer and are called **"moments of pain"** when the impact is negative. Finding these "moments" of the client is relevant for companies if they want to impact on their perception and create a remembering experience throughout their journey as a consumer.

As can be seen from the above definitions, the experience is intimately linked to customer expectations and brand promise, so the perception is different for each sector and for each brand. And it is this attribute of temporality that gives the interaction its unique character and so the brand must exercise extreme control over the entire interaction (Blut, Beatty, Evanschitzky & Brock, 2014).

However, different moments can be found in the client's journey, as a function of the circumstances of the interaction. Thus, according to Cohen (2013), there are four different moments:

The **zero** moment of truth occurs when the client recognises the need and seeks to satisfy it. Surprising the customer with an excellent information architecture is key to their experience.

At the **first** moment of truth, the customer knows all the information and weighs the purchase decision against real-life information. This moment is crucial since the acquisition or use of the good or service will depend on it.

At the **second** moment of truth, the customer consumes the product or service. He can finally confirm whether it meets his needs and confirms his expectations. Exceeding them will be a determining factor in customer retention and a background to the recommendation.

At the **third** moment of truth, the client feels bound to the company and wants to be part of it. Maintaining a differential contact with the customer is necessary for generating a reciprocal relationship and closing the link between the user and the brand. Clearly, this third moment is the key to spontaneous recommendations (Reichheld, 2014), that are more shocking and sincere than the traditional ones, sometimes the result of pressure groups (Bilgihan, Barreda, Okumus, & Nusair, 2016).

As business researchers, the company must understand each moment of truth in order to discover the determinants and consequences of each moment. To do this, the company will use Applied Data Science to collect as much data as possible about the transaction. One of the most conclusive of these is the Voice of the Customer, not only because it shows the opinion about the interaction, both on the quality and on the way of offering it, but also because it allows the company to gather its vision of the improvement, to infer its emotions and to strengthen the bond that unites them.

Strategic Importance of the Customer Experience: Consequences for Retention and Sustainable Growth

Offering value to the customer through consumer experiences generates **satisfaction** (Ban & Kim, 2019). This evolution of the relationship with the customer from a specific transaction to a stable, differential

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

and satisfactory relationship, which is achieved from an excellent customer journey, provides a **loyal clientele** that seeks differential experiences (Kotler, Armstrong, Cámara and Cruz, 2004; Rodríguez-Ardura, Maraver-Tarifa, Jiménez-Zarco, Martínez-Argüelles and Ammetller, 2018), in other words, delight in consumption (Ali, Kim, Li and Jeon, 2018).

The consequences of maintaining an experiential relationship over time are immediate:

- It allows us to **know** the customer, comparing the consumer experience received with their expectations (Pine & Gilmore, 1999; Bilgihan, Kandampully & Zhang, 2015), and illuminating the relationship between the experience and the final satisfaction obtained (Prebensen, Kim & Uysal, 2015). This knowledge is vital for the improvement of the product/ service and how to offer it in an attractive way to the customer.
- It consolidates **trust** in the brand as the client perceives the quality offered (Skålén, Gummerus, Von Koskull and Magnusson (2015) and the interest in their well-being and commitment to the relationship (Morgan and Hunt, 1994).
- Promotes the user's intention to **recommend** the company (Reichheld, 2014).

In the case of digital relationships, an excellent customer experience is also warranted:

- The intention to **recommend** the digital channel (Lorenzo Romero, Alarcón de Amo & Gómez Borja, 2011).
- The intention to **repeat** the use of the channel (Chen & Phou, 2013).

Translated into business KPIs (Key Performance Indicator), an excellent customer experience means:

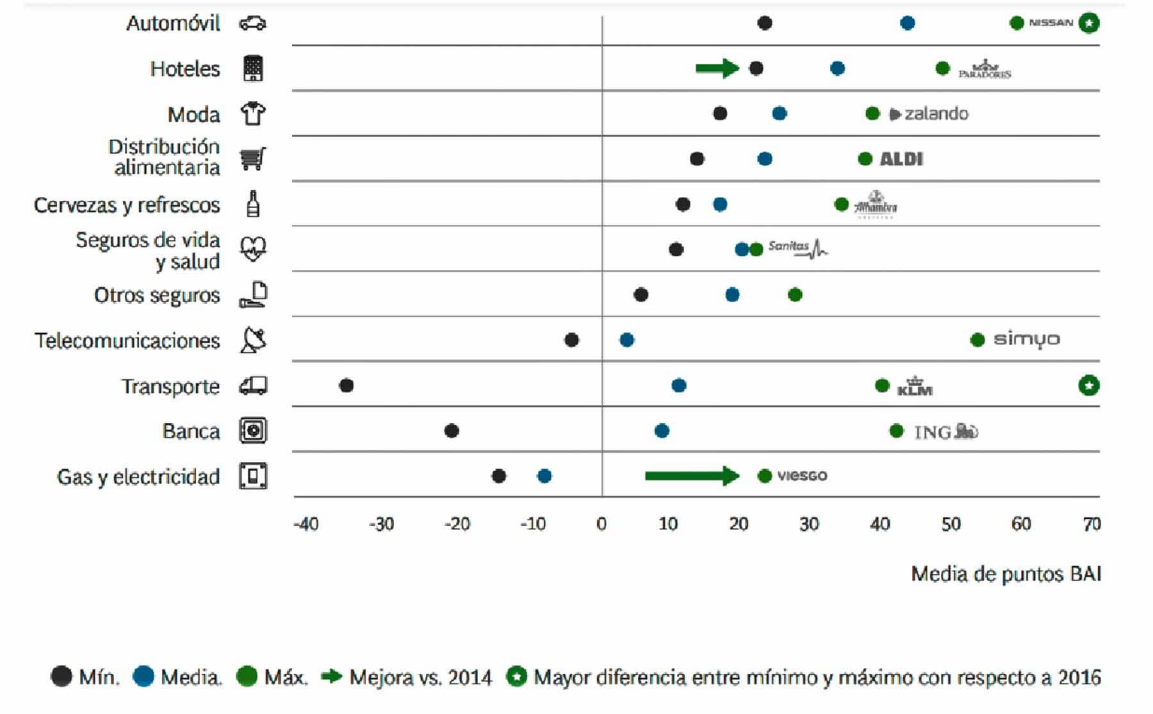
1. **Increases in sales and customer/user loyalty**: According to the latest Forrester (2020) report, companies working on customer experience management experienced 14% higher sales than those without such initiatives
2. **Increases in new contracts.** Consumer satisfaction encourages recommendation and is an opportunity for companies to have a digital presence, so that recommendation can become a trend thanks to social networks and influencers (iabSpain, 2018). The impact is notable: As can be seen in the case of Bankia (Sabio & DEC, 2020), it can mean up to 63% new registrations (Medallia & DEC, 2020), an increase in the satisfaction index of 20% over 6 years, and an increase in recommendations up to 92% (Sabio & DEC, 2020). In short, companies that invest in customer experience manage to double their real and spontaneous recommendations compared to those that do not, as can be seen in the Boston Consulting Group's report on 2014-2018 comparisons (BCG & DEC, 2018) (see Figure 3).
3. **Improving the image and reputation of the brand and reducing the cost of marketing campaigns.** The customer experience improves the customer's perception of your brand, so that it is not necessary to maintain the memory of the brand promise through advertising techniques (Sanny, Arina, Maulidya & Pertiwi, 2020).
4. **Reduction of complaints or claims.** An excellent experience reduces user discomfort and complaints, so dealing with and reducing customer complaints not only helps to stop the loss of customers, but also increases the ROI of the cost of the complaint (Ntwali, Kituyi & Kengere, 2020). Complaints are not problems but opportunities to grow and improve, so any investment is profitable as the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CepymeNews report (2020) indicates. What's more, facilitating complaints can improve brand perception. One of Bill Gates' most famous quotes is "Your most dissatisfied customers are your greatest source of learning", so understanding complaints helps build brand loyalty, improving the consumer experience.

5. **Resistance to change**. Consumers tend to be conservative as change means risk (Zhang & Li, 2019; CepymeNews, 2020). And nothing better than a satisfied customer to make the change unjustified.
6. **Lower price sensitivity**. The Customer Experience is responsible for more than two-thirds of customer loyalty, above and beyond the brand and price effect (Gatner, 2019).
7. **Possibility of greater price benefit in price**. In line with the previous reference, 42% of consumers worldwide would pay more for a "friendly and welcoming experience" (PWC, 2020).
8. **Better use of resources and reduction of personnel costs.** Thanks to Artificial Intelligence, it is possible to automatize processes that are or have been a source of satisfaction for the client or that help to achieve it. Being able to industrialise and to automate certain aspects of the experience allows for a better use of the available personnel resources, which can be fully dedicated to the emotional care of the client (Temkin, 2019).
9. **Better work atmosphere and identification of personnel**. To be happy at work is difficult if the employee's mission is to attend to dissatisfied customers. Therefore, improving the customer experience improves the working climate by reducing worker stress, wear and tear and the pressure on the worker. In addition, the customer experience must be an integral vocation of the company,

Figure 3.
Source: (Report on Customer Experience in Spain. Comparison 2014-2018, BCG & DEC, 2018)



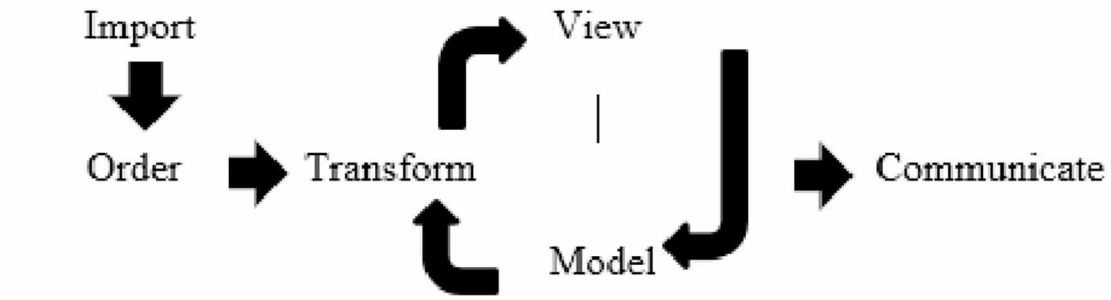
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

where all the departments or people involved in the different points of the customer's journey must work as a team to achieve an overall satisfactory experience. This promotes cohesion and identification with the company's values, a common objective, identification of the staff with the brand and commitment to the same goals (Izo, 2019b).

THE APPLIED DATA SCIENCE AND ITS CONNECTION TO THE CUSTOMER EXPERIENCE

Understanding the customer is basic to the success of companies, but as we have already argued, the large **volumes** of data generated by the user experience make the process impossible in **real time** by traditional means and generate the most appropriate response to the customer in real time. This is how the application of Science in business was born, the objective of which is to **collect, process and elaborate** useful information for customer-centred business management (see figure 4). In this way, not only do new products arise such as intelligent environments (connected homes, assisted driving, usage fees, etc.), but it also allows new forms of communication as another form of experience. Thus, thanks to technology and computer knowledge, it is possible to provide a comprehensive service where the omni-channel customer experience provides a subtle, warm and differential treatment that humanises each interaction, surpassing the effectiveness of traditional means.

Figure 4.
Source: (Vázquez Brust, 2020)



By way of introduction, it can be said that Applied Data Science consists of applying programming techniques to the data collected, which requires the development of skills in four areas:

- **Programming.** Programming commands need to be entered to explain to the analytical computer tools what is required of them. That is to say, we must descend to programming languages that reduce complex tasks in steps to follow in the consequence of results. These new high-level languages require a high level of sociological and marketing knowledge as the key to this science is not only knowing what data to collect but also what value they have for the company.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Statistics.** Thanks to technology it is possible to predict behaviour models from the reliable underlying relationships between variables, through more than 10,000 iterations of random subsamples, which allow us to confirm hypotheses and establish robust variables from the attributes that define them (Hair, Hult, Ringle & Sarstedt, 2017).
- **Domain knowledge.** We are all clients or companies in our daily performance: user, patient, student, client, public or private entity or organisation. Domain knowledge allows us to understand if the statistical answers obtained are coherent and provide value and new knowledge.
- **Communication.** This interdisciplinary collaboration between Science and Business requires adequate communication that can be translated into timely business actions from the new knowledge. As indicated in the **Cluetrain** manifesto (Cluetrain, 1999) "In just a few years, the current homogenised "voice" of the business world -- the sound of corporate missions and official brochures -- will seem as far-fetched and artificial as the language of the French court in the 18th century". ADS and AI make it possible to find this new style of corporate and customer communication adapted to each customer profile. What's more, the new Iot devices allow for customisation by each user as a further attraction of the product. In this line, it is expected that by 2030, 81% of consumers will expect to interact with chatbots and 8 out of 10 consumers of smart assistants will make their purchases online (PuroMarketing, 2019).

As can be seen from the exhibition, ADS is much more than just collecting information and applying some statistical correlation. Its strategic interest has made it an obligatory reference topic on the web, with more than a billion entries for the Applied Data Science category (the latest reference on Google for October 2020) and a new **discipline** of study in universities.

In the following sections, we will describe the scope of Applied Data Science, the types of data analysed, its applications in Industry and Business, and how the evolution of correlations has led to new applications such as Artificial Intelligence and Blockchain.

Why Companies Should Measure

In this new socio-economic context of the customer experience, ADS and in particular AI are presented as a great opportunity for brands as they allow them to process information about the customer in real time and personalise each interaction (Torrent-Sellens, 2020). Companies worldwide have understood this vision of the need for innovation and are investing in consolidating customer data platforms (53%) and real-time decision engines (45%) (CepymeNews, 2020). The implementation of artificial intelligence and automatic learning has become a priority in the strategy of customer experience departments, where 36% of the companies consulted are already using AI and 28% have planned to invest in them (Adobe, 2020).

In each interaction, all data from the customer is relevant. Thus, we have data:

- **Structured,** where its content can be delimited in lower units of information, such as filiation data, purchase data, payment data, etc.
- **Unstructured,** such as texts in comments on social networks or other media, body posture, facial expressions, physiological response, dispositions to action, actions or personal judgments (Pelegrín Borondo, 2013). From this multiplicity of sources, companies can access highly relevant customer information: their emotions (Lindstrom, 2019; Lourido, 2017; Pelegrín-Borondo et al., 2017).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

As Table 2 indicates, companies can turn **feelings** into data by discerning patterns of behavior (obtained from ADS and IA) that reveal multiple emotions. Such is the case of anger, sadness, euphoria, satisfaction, surprise, enthusiasm or apathy, are determining factors in people's behavior (García Haro, 2018). With this information, the company can glean customers' feelings and preferences and determines, tone and style of communication, personalisation of products and services and the most appropriate channels in which to serve them.

Emotions, defined by Izard (1977:48) as "a complex process with neurophysiological, neuromuscular, and phenomenal aspects", are the object of multiple researches in the theoretical framework. As indicated by Pelegrin Borondo (2013), citing Kleinginna and Kleinginna (1981), up to 101 definitions or the 175 different synonyms for emotions can be considered, counted by Cohen, Pham and Andrade (2006). Regardless of the specific form of expression, it is important to differentiate between affection, mood and emotions, a common mistake in the business world. Thus, emotion is temporary, of greater intensity and feeling than mood and of a less complex nature. Moreover, the latter can harbour various emotions. And, although they are different concepts, feelings have been equated to emotions by some authors (Martínez Sánchez, 2008).

Table 2. Basic emotions identified in the research framework

REFERENCE	BASIC EMOTIONS
Arnold (1960)	Anger, aversion, courage, dejection, desire, contempt, fear, hate, hope, love and sadness
Ekman, Friesen & Ellsworth (1982)	Anger, displeasure, fear, joy, sadness, surprise
Frijda (1986)	Wish, content, interest, surprise, wonder
Gray (1982)	Anger, terror, anxiety, joy
Izard (1971)	Anger, happiness, disgust, anxiety, fear, guilt, interest, joy, shame, surprise
James (1884)	Fear, regret, love and anger
McDougall (1926)	Anger, disgust, joy, fear, submission, tenderness, wonder
Mowrer (1960)	Panic, pleasure
Oatley & Johnson-Laird (1987)	Anger, displeasure, anxiety, happiness, sadness
Pankesepp (1982)	Expectancy, fear, anger, panic
Plutchink (1980)	Acceptance, fear, surprise, sadness, aversion, anger, anticipation and joy
Tomkins (1984)	Anger, interest, joy, displeasure, anguish, fear, joy, shame, surprise
Watson (1930)	Fear, love, anger
Weitner & Graham (1984)	Happiness and sadness

Source: (Ortony & Turner, 1990 in Pelegrín Borondo, 2013)

In short, subjective evaluative judgement or emotional state, distinguishing the client's emotions and moods in the unstructured data collected is a complex task since the difficulty of the source is compounded by the psychographic and socio-cultural difference of each individual. However, thanks to IA, it is possible to distinguish the generating or triggering stimulus (Russell, 2003) from the emotions since they do not all have the same intensity (Frijda, 1986) or relevance for the client (Martínez Sánchez, 2008), and lets to decide whether it is necessary to produce an immediate response (Scherer, 2005). But

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in any case, the reaction obtained will be a source of new knowledge for their emotional architecture and future behaviour. It is worth mentioning the following tools that allow the underlying emotions to be found in the Voice of the Client:

- **Text Mining**: it allows to know the mood and emotions of the client under statistical patterns of semantic meanings according to the use of negations, adjectives, detected verbs and other data of the text.
- **Speech Analytics**: similar to Text Mining, it allows for lower emotions of the client from the tone of voice and the type of verbal language used.
- **Neuromarketing**: this discipline also collects customer data from biometric recognition. This is the case with changes in temperature (tomography); eye movement; pupil width; keyboard pressure; electrical changes in the brain (encephalography) and magnetic changes (magnetoencephalography); etc.

Types of Customer Data: Structured and Unstructured and Irrational

Collecting data, as has been argued, is essential for implementing strategies based on ADS and AI, but not all data are valid. The information collected must be **relevant** and provide **knowledge** to the organization in the context of the business. Companies that want to focus on their customer and develop customer experience strategies need a program of VoC, that can listen to, measure, analyse, and act in real time on collected data and feedback. and act in real time on customer feedback and data collected. Thanks to the **Enterprise Feedback Management Systems** (EFMS), this data collected must guarantee a predictable, objective and reliable result in the management of the customer experience.

It is therefore necessary to capture the opinion of customers through traditional means, the new digital ones (**web, email, QR code, SMS, social networks, fixed and mobile devices, applications, etc.**) and those generated by the environment (e.g., **geolocation and the Internet of Things**). Because these data are very diverse, they must be correlated with the customer and the specific consumption environment in order to decide whether to generate, modify and monitor individual customer behaviour in real time. And all this is based on the personalisation that can be deduced from knowledge of previous actions.

There are multiple tools that allow us to explore and analyse the data. Our human condition is not always coherent given our irrational aspect (limbic system), far from our rational cognitive system. For example, choosing a brand that transgresses the customer's tastes, or making compulsive purchases, can be symptoms of sadness or guilt. Anticipating situations like this are beyond any current technology, but they allow us to create patterns of behaviour to be exploited in the future.

Such complex information evolves over time, so the concept of omnichannel, use of traditional and digital channels in communication with the client, must be guaranteed. In the new digital era, it is essential to collect data on the digital customer experience if companies want it to be a preferential channel for their users, with the advantages that this entails for the company, such as continuous contact and direct personalisation of the offer (Xiao, Zhao, Pan, Song, Zheng & Yang, 2019)). Thus, in a web context, where self-service must be guaranteed, thanks to the information on the behaviour and feelings of its customers, the company can provide an **intimate, personal and close atmosphere** with its users that makes the presence of an agent unnecessary (Palm, 2016), so that the feeling of trust and security is not lost in a non-human context. Advanced technologies with high multimedia capabilities, depending on the context, segment and customer preferences, can offer and generate surprising experiences that exceed

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

any expectation. Moreover, given the relevance of customer feedback, the survey itself has become yet another surprising and visual experience (Wise & DEC, 2020) with a very clear underlying message: the company's concern for customer welfare and satisfaction.

One of the big problems with consumer behaviour research is that scales and forms may not reflect the true feelings and experience of the customer so that a data obtained, if true, may not necessarily reflect the reality of the customer but only appropriate and honest answers to poorly formulated questions. Therefore, this data in structured or unstructured format must be debugged to avoid:

- **Duplicity**
- **Uncoordinated**
- **Lack of focus on the objective**
- **Lack of action orientation**

The valuable information obtained from unsolicited feedback data is relevant, such as social networks, chat, contact centres, Customer Relationship Management (hereinafter CRM), and open comments, which normally offer unstructured, unformatted data. This type of channel requires advanced text processing and analysis using natural language and automatic learning technologies to provide a complete picture of the customer's life.

Business and Customer Oriented Data

In order to act on the customer, the first step is to measure their experience as a brand user. Each customer contact point with the company generates a significant amount of data from multiple sources and formats and from various channels (Stiga, 2017 in DEC, 2017). This complex and massive information, sometimes totally unstructured, requires computer processing to identify and consolidate the data collected. Once ADS transforms the data into relevant information about the interaction, it is possible to understand the underlying relationships of the identified variables and detect patterns of behaviour. Thanks to these psychographic segmentations of the customer, the company can design marketing actions in a reactive and proactive manner to improve both the experience of its customers and that of its employees.

This invaluable information is exploited for the business world under multiple perspectives in order to industrialise the experience (Torrent-Sellens, 2020). In this way, we can exploit:

1. **Descriptive analytics,** which indicate what has happened or is happening in the company and with the client. Also called Business Intelligence, the reports they produce inform the company's directors about the positive aspects of the situation being analysed.
2. **Predictive analytics**, which shows what is going to happen and why. That is, it measures the probability of "something" happening. In this case, the information is structured into business rules, propensities and behaviour patterns. In this way, it is feasible for the company to observe before acting.

Finally, the stored data must be transformed into indicators for the ADS and for the business, in order to be metadata for other analyses. In other words, it must be possible to deduce from the data customer indicators, the nature of which must be:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Relevant:** it must bring knowledge to the organization.
- **Predictable:** the indicator must be consistent and ascend or descend according to the sign of experience.
- **Objective:** the indicator must have a clear meaning for the organization of the client's representativeness.
- **Calibration:** it means the same in every organisation and environment given the global nature of the economy.
- **Reliable:** the data meets the customer's expectations of security and trust.
- **Simple**: it does not respond to complex nature in such a way that it can be variable and interpreted by higher level relationship models. This level of complexity can be achieved through SEM models and latent variables where the appropriate formative and explanatory latent relationships will be analysed.
- **Transparent:** collected in accordance with current legislation.
- **Consistent:** obtaining the data must not contaminate it, avoiding any bias in its handling.

The data can show relationships between segments (cluster analysis); explain relationships between variables (principal component analysis); show indicators that define a variable (factor analysis); their predictive ability (Stone Geisser index, Ning & Liu, 2008); confirm the relationship between variables; reduce bias in estimates; and behaviour over time (time series), among other results.

According to the Capgemini report (2020), the big challenges in ADS and business-oriented AI allow companies to achieve:

- The **hyper-personalisation** of products and services.
- **Multichannel** data collection.
- The **continuity** of information throughout the **customer journey**.
- **Transparency** in the processing of customer data, which must comply with current **legislation** in accordance with the ethics and security of the customer's confidential information (ISO 25010, 2020).
- **The traceability of the data** so that it can fully inform the consumer context.
- It must have **value** for the company in its ability to predict results.

Evolution of Indicators: From Descriptive Statistics to Big Data and Small Data

Although one of the traditional forms of exploitation has been from information coming from Big Data, it is in Small Data where spectacular results are being produced for the design of personal experiences in the mentioned sensitive field. But what are the differences? In Big Data, the 3V criterion is met, which is a large volume of information, high processing speed, and a variety of information from clients, channels and processes. Its application is immediate for strategic purposes, for marketing decisions on product and price and for logistics in operations and campaigns. That is, huge amounts of correlated numerical data, impossible to process by traditional metrics and methods.

But it is not only the volume that is its greatest feature. Really, there are 6 Vs that define the Big Data technique:

- **Volume**: capacity to process huge amounts of information.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Speed**: capacity to respond in real time.
- **Variety:** ability to process data from all sources and of any nature.
- **Veracity:** and respond to the reality of the clients.
- **Feasibility:** it must be possible to process it.
- **Value**: they must have value for the company.

On the other hand, S**mall Data**, reflects aspects of customers that are impossible to quantify from digital environments, such as the customer's lifestyle, their opinions and the emotions they feel when consuming the brand. These perceptions are gathered from qualitative techniques, such as direct observation and in-depth interviews (Taylor & Bodgan, 1984). This technique is the source of the shadowing methodology (DEC, 2017) or day in the life of the client, which allows for a global vision of the consumer context: culture, sociology, ethics, trends, values, etc. that undoubtedly influence the client's future decisions.

Small Data reflects the idea that measuring the Customer Experience is about trying to discover the emotional stories behind the numbers (Lindstrom, 2016). In this way, it is possible to draw the **map of empathy** (Guimerá, 2015, quoting Dave Gray, founder of Xplaner and author of the methodology), which shows the customer's holistic vision: their expectations, needs, and unique worldview in a way that reflects their concrete reality (Lindstrom, 2016, 2019).

Applying AI to these information sources not only facilitates a deeper and closer approach with a company's users, understanding what motivates or frustrates them, but is also the basis for designing hyper-customized products and services (Lindstrom, 2016). With Small Data, one can directly and qualitatively establish the **stakeholder map**, reflecting a company's actual users and target audience, and thereby understand segments as actual human beings (DEC, 2017).

In short, Big Data represents the macro vision while Small Data is the micro vision, but both are necessary for studying customers' needs in the context of the market.

Artificial Intelligence and Blockchain

Notably, these generated behavioural profiles become **metadata** for ADS, which can be consolidated into **behavioural patterns.** Artificial intelligence, created as a discipline by scientist John McCarthy (Dartmouth Conference in 1956), allows to make its own decisions on the basis of acquired knowledge and its ability to predict results.

Predictive behaviour models consolidate relationships that allow the construction of increasingly complex variables. In this way, talking about customer expectations in the digital area, for example, means actually referring to reputation, social pressure, technology adoption, perceived utility and ease of use (Méndez Aparicio, 2019). This allows us to directly relate this customer experience concept with others, without having to reinterpret all the **underlying relationships,** enriching the stored knowledge (Shiau, Sarstedt & Hair, 2019). Thus, similar to the functioning of our brain, the units of information are connected to each other, creating true **networks of information**, about the attributes of the product or service, associated to the experience lived in the interaction. In this way and similar to human cognitive processes, the IA collects and processes qualitative and quantitative information from the customer in their interaction with the brand. Through complex computer processes, AI is able to recreate the set of emotions, sensations and sensory images felt by the individual associated with the attributes of the product or service (Walls, et al., 2011; Duerig, Melton, & Stöckel, 2013), storing this information in **virtual neural networks** (Gutmann, & Konagaya, 2019) as if it were human memory (Sierra Diez,

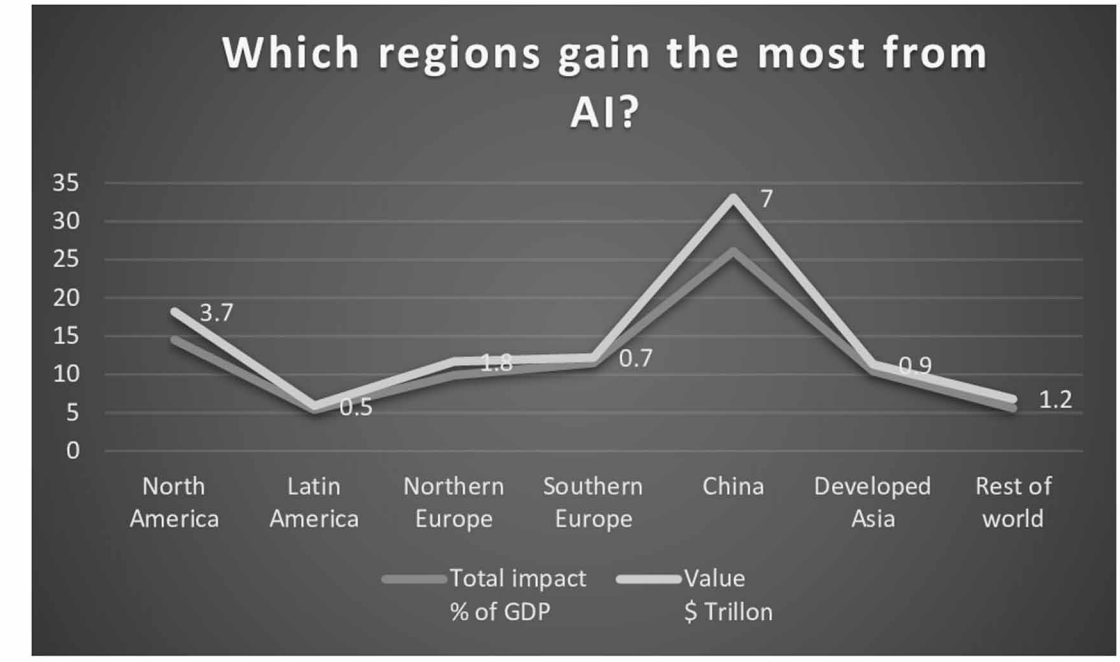
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Froufe Torres, & Falces Delgado, 2010; DEC, 2017). Thus, information can be obtained on how a positive experience generates a positive response in the user towards the brand and predicts his behaviour and future interaction as a customer. This new cognitive science allows humans to explore our perception and reasoning processes as well as replicate consistent replicate consistent search algorithms (Lee, Singh & Azamfar, 2019). This reinterpretation of reality requires superior intelligence from technology capable of understanding users even in non-rational behaviours, such as digital behaviour (Liuqu, Fan & Fu, 2015) where, due to the perceived risk inherent in the uncertain nature of online shopping and services, online customers tend to be less rational.

The relevance of artificial intelligence in industry and business is already a reality. According to the PWC report (PWC, 2017a), the impact of AI on the world economy will be more than 15.7 trillion dollars by 2030, equivalent to a 14% growth in world GDP, as shown in Figure 5. According to the IZO Consultancy (2019a) the future of customer service lies in the automation and personalization of experiences that can produce greater customer engagement and effective satisfaction of their needs.

One of the most relevant applications of AI at the customer level is the **chatbot**: software units designed to run automatically and autonomously. For example, chatbots can acting on the calendar, restaurant reservations, displaying information associated with a keyword, placing food orders, etc., simple operations that can be linked. One of the most visible applications is the chatbot, which allows a conversation with the customer without human intervention.

Figure 5.
Source: (own elaboration adapted from PWC, 2017b)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Another of the most recent applications is its combination with the new **blockchain** technology, which according to Gartner (2020) will contribute up to 2030 3.1 trillion dollars to the world economy. Blockchain stores data in a **distributed, decentralized, and immutable way**. Its combination with AI allows instant access to information, taking a giant step in the use of data. Thus, it is possible to create decentralized data sets with authorized access by trusted third parties, promotes security in its treatment, allows monetization without risk of breaking the mutual trust and facilitates business decisions in a global economy (BBVA, 2019). This combination allows for an exponential escalation of mutual benefits where encryption algorithms are improved and the power of IA is increased (BBVA, 2019).

This huge amount of information provided by the customer's **customer journey** and the psychographic profile of their environment allows us to predict future consumer behaviour from their inference in contextualized consumption models. There are multiple applications, we can cite the most important among them:

- **Automating** certain decisions that optimize the process in time and resources from relevant customer information based on specific criteria of the process that is happening. For example, intelligent CRMs, which cover the entire customer lifecycle.
- **Biometric recognition** of the client to authorize processes and/or detect emotions. For example, facial recognition.
- **Customize** the price according to the use in time and/or location of the client. For example, the car connected in insurance, which detects the real risk and therefore adjusts the rate.
- **Adapt** the product or service in real time. For example, Zara's smart changing rooms, which allow you to visualize the effect of a virtual garment.
- **Facilitate the interaction** with the agent and even allow the customer to be satisfactorily self-sufficient. For example, in the digital channel, through the "clicks" made (Google Analytics) or in a face-to-face environment through the mobile (geolocation), it is possible to obtain customer navigation patterns in the environment and to obtain information. Through Automatic Learning (or machine learning, BBVA, 2019) it is possible to infer patterns of behaviour and guide the customer to his destination.
- **Facilitate purchases** according to previous patterns with the incorporation of updated information. For example, Amazon.
- **Multi-channel automatic learning** (or machine learning), which allows for the detection of key words with relevant meaning for the customer (natural language comprehension, NLC.

One of the most important applications of AI is natural language processing (NLP), applicable in the mass processing of unstructured data. Thanks to sophisticated algorithms it is possible to create a new internal vocabulary by means of words grouped by proximity according to patterns provided by qualitative research (focus groups, participant observation, in-depth interviews) or other methodologies (Balyan, McCarthy & McNamara, 2020). Each successful catalogued finding is considered a training that feeds back into the information architecture.

In short, thanks to AI, it is possible to interpret behaviours according to previous consumption patterns, which allows the generation of new and greater knowledge regarding future customer behaviours.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

BUSINESS SOLUTIONS AND RECOMMENDATIONS

As developed in the previous sections, the Applied Data Science and Artificial Intelligence open up a world of infinite possibilities for marketing practices. Of the many combinations of technology, business and context data, the following business implications are worth mentioning:

1. Facilitates **campaigns to attract** new clients by having precise information on each segment (Vargha, 2018).
2. Allows for the **collection and processing of the client's voice,** essential information for adapting products and services to the needs of each consumer (Deloitte, 2018).
3. Allows for a **continuous dialogue** with the client in real time, promoting digital humanism as a user experience (Martín Lineros, 2020).
4. Establishes and confirms **predictive models** on customer additions and deletions, allowing the company to take the appropriate measures to attract and retain customers (Méndez-Aparicio, Jiménez-Zarco, Izquierdo-Yusta & Martinez-Ruíz, 2020).
5. Provides customer experiences that encourage their **recommendation to third parties,** ensuring sustainable growth of the company (Lin & Lekhawipat, 2014; Reichheld, 2014)
6. Guarantees reliable company **budgets** by having the ability to extrapolate results from the comparison of final benefits obtained and customer behaviour throughout the customer journey (DEC, 2017).
7. Facilitates the **improvement of the ROI** of the digital transformation by designing the best products and tools for the customer (Turletti, 2019).
8. Visibly improves the company's **brand image** through process innovation, differentiation and quality (Quevedo, de la Fuente & Delgado, 2005).
9. Facilitates and optimizes **advertising campaigns**. According to historical data from previous campaigns, statistical reports and intelligent recognition of relevant words with excellent return on investment (ROI), it is possible to measure the scope and investment needed in new campaigns towards the client (Wierenga, 2010).
10. It allows to analyse the expected **customer ROI,** both in terms of acquisition and loyalty, and to show the investment limit in the customer experience for each consumer according to the expected return on investment (DEC, 2017).
11. Retains **unsatisfied customers** through concrete and effective actions against the dissatisfied customer ("close the loop"), anticipating solutions and surprising them from a deep understanding of their expectations (Qualtrics, 2020).
12. Allows the correlation of customer and employee experience with **business indicators** (CepymeNews, 2020).
13. It facilitates the management of **distribution and storage logistics,** essential in the explosion of e-commerce (Pandian, 2019).

As can be seen from this section, as well as the above-mentioned business-oriented results, the applications of Applied Data Science and artificial intelligence are enormous and only possible from the **scientific treatment of information**. As a guarantor of experience and as a control and management of the expense and sustainability of the companies.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Finally, it is worth mentioning the relevance for companies to comply with current legislation. In 2018 the General Data Protection Regulation of the European Union (hereinafter **GDPR**) was approved, a law that was passed in April 2016 and came into effect in May 2018 (Intersoft Consulting, 2020). This law replaces the previous Law on Protection of Personal Data and guarantee of digital rights (hereinafter LOPD) (BOE, 2018) and absorbs the Robinson list, that regulated the sending of advertising to customers. In this law, it is regulated at European level, by which channel companies should contact customers, for what, what data can be requested that are relevant to the business and how long they can be stored. Most importantly, companies must maintain traceability of consent throughout the *customer journey*, which can be a real problem in mass data processing (Zarsky, 2016).

FUTURE RESEARCH DIRECTION

Companies are immersed in a profound digital transformation. Although it was incipient and pioneering at the end of the 20th century, adapting to the connected society offered by the Internet, the advances in technology have advanced industries and businesses. Thus, in the midst of the 21st century, it can be said that it is now that companies have understood and engaged in a profound change to make possible the collection and processing of information, multi-channel adaptation and adoption of new solutions for the Internet of things and mobility (Cisco, 2020). The so-called digital transformation has only just begun (Vial, 2019). **Marketing 5.0 is an immediate reality**. The irruption of 5G is changing way of communicating (Martin Lineros, 2020). Third-party certification (blockchains) is already a reality that transcends traditional controls (Pilkington, 2016). This has led to a remodelling of all information systems, measurement and change in processes (Elsden, et al., 2018).

In this change of strategy of the companies, to the mentioned **customer experience** must be added the **employee experience**, who must adapt to new processes of his company (DEC & Bain & Co, 2020), but also to the demands of the **new consumer**, empowered in his needs (Auh, Menguc, Katsikeas & Jung, 2019). As these authors indicate, these changes are requiring a profound revision also of the business structures.

Thus, the **artificial intelligence**, created as a discipline by the scientist John McCarthy (Dartmouth Conference in 1956) comes to stay in the organizations. The big companies consolidated in the previous century, see how their strategies are becoming obsolete because of their difficulty to react to the change and the innovative incentives take over. In this new business environment, characterized by its volatility, uncertainty, complexity and ambiguity (Oliva & Kotabe, 2019), the need for client-company knowledge is vital, where **innovation** and **business know-how** come together. This formula has already been adopted in new solutions by large corporations that see the alliance with **start-ups** as a way to test their initiatives as incubators for their projects and to launch pilots of new products and processes (Lange & Johnston, 2020).

This new business vision allows the growth and coexistence of multiple initiatives including co-creation with the client (Mendez-Aparicio, et al. 2020) so that consumers are part of the new business solutions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Research on Applied Data Science and Artificial Intelligence in Business and Industry develops with due depth how these new disciplines improve the business from a new policy focused on its customers.

Through the different sections of this chapter, you can appreciate how technology accompanies the customer throughout its lifecycle, generating valuable information for business and industry. From the pre-sale, offering the prescriber privileged information, to the experience of surprising experiences at the time of purchase or consumption and customer service in the after-sales, **Applied Data Science improves satisfaction** in every interaction with the customer. Thanks to the new tools, a **new form of communication is possible, continuous, warm and close,** which allows the link to be maintained throughout its life cycle in the company. This new vision of the customer evolves the concept of customer journey towards the vision of the empathy map, where the feelings and emotions of the customer are collected.

This intangible and powerful value allows companies to accompany the customer, remain in his memory and be his best option in the decision processes (Gómez Borja, 2014). In a very complex moment of the economy, global and shaken by the pandemic situation that affects us, being a guarantor of the physical and emotional needs of the consumer is changing the concept of marketing of companies. Knowing the client, lifestyle, context and groups of influence is only possible through the combination of information provided by Big Data, Small Data and Artificial Intelligence. In a connected society, **information is a new tangible value that allows brands great business opportunities**, but also increases the risk for them, since they must permanently loyalty their customers (DEC Congress, 2020) and anticipate the evolution of consumer expectations. And this perspective is only possible from the deep knowledge that allows the treatment of information. As it has been indicated in the introduction, it is generated in the network everyday hundreds of Teras of information that it is only possible to understand, to analyse and to consolidate as knowledge from this new discipline as it is the Applied Data Science.

REFERENCES

Adhikari, A., & Bhattacharya, S. (2016). Appraisal of literature on customer experience in tourism sector: Review and framework. *Current Issues in Tourism*, *19*(4), 296–321. doi:10.1080/13683500.2015.1082538

Adobe. (2020). *Digital Trends Study: 2020 Is The Year Of CX-Centric Business Transformation*. Retrieved January 15, 2021, from https://cmo.adobe.com/articles/2020/2/digital-trends-study--2020-is-the-year-of-cx-centric-business-tr.html#gs.imxcnu

Ali, F., Kim, W. G., Li, J., & Jeon, H. M. (2018). Make it delightful: Customers' experience, satisfaction and loyalty in Malaysian theme parks. *Journal of Destination Marketing & Management*, *7*, 1–11. doi:10.1016/j.jdmm.2016.05.003

Alonso-Cortés, F. (2020*). ¿Hacia dónde vamos en la nueva era del cliente?* Retrieved December 16, 2020, from https://www.marketingdirecto.com/especiales/the-future-of-advertising-especiales/hacia-donde-vamos-en-la-nueva-era-del-cliente

Aoki, N. (2020). The importance of the assurance that "humans are still in the decision loop" for public trust in artificial intelligence: Evidence from an online experiment. *Computers in Human Behavior*, 106572.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Asociación para el Desarrollo de la Experiencia de Cliente. (2017). La experiencia de cliente rentable. Libro colaborativo. Ed. DEC, Madrid 2017.

Asociación para el Desarrollo de la Experiencia de Cliente. (2019). *Behavioral Economics. VI Congreso Internacional sobre Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/eventos/congreso/congreso-2019/

Asociación para el Desarrollo de la Experiencia de Cliente. (2020). *The Experience Revolution. VII Congreso Internacional sobre Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/eventos/congreso/congreso-2020/

Auh, S., Menguc, B., Katsikeas, C. S., & Jung, Y. S. (2019). When Does Customer Participation Matter? An Empirical Investigation of the Role of Customer Empowerment in the Customer Participation–Performance Link. *JMR, Journal of Marketing Research*, *56*(6), 1012–1033. doi:10.1177/0022243719866408

Bain & Co, & Asociación para el Desarrollo de la Experiencia de Cliente. (2020). *V Informe de Madurez de Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/presentacion-de-informes/v-informe-de-madurez-de-experiencia-de-cliente/45773/

Balaguer. (2020). *¿Es posible analizar el ROI en la experiencia del cliente?* Retrieved March 15, 2021, from https://micliente.net/roi-en-la-experiencia-del-cliente/

Balyan, R., McCarthy, K. S., & McNamara, D. S. (2020). Applying natural language processing and hierarchical machine learning approaches to text difficulty classification. *International Journal of Artificial Intelligence in Education*, *30*(3), 1–34. doi:10.100740593-020-00201-7

Ban, H. J., & Kim, H. S. (2019). Understanding customer experience and satisfaction through airline passengers' online review. *Sustainability*, *11*(15), 4066. doi:10.3390u11154066

BBVA. (2019). *Machine learning': ¿qué es y cómo funciona?* Retrieved November 25, 2020, from https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/

BBVA. (2020). *Blockchain y la inteligencia artificial: ¿La pareja ideal?* Retrieved December 23, 2020, from https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/blockchain-y-la-inteligencia-artificial-la-pareja-idealblockchain-y-la-inteligencia-artificial-la-pareja-ideal/

Bilgihan, A., Barreda, A., Okumus, F., & Nusair, K. (2016). Consumer perception of knowledge sharing in travel related online social networks. *Tourism Management*, *52*, 287–296. doi:10.1016/j.tourman.2015.07.002

Bilgihan, A., Kandampully, J., & Zhang, T. (2015). Towards a unified customer experience in online shopping environments. Antecedents and outcomes. *International Journal of Quality and Service Sciences*, *8*(1), 102-119. doi:10.1108/IJQSS-07-2015-0054

Blut, M., Beatty, S. E., Evanschitzky, H., & Brock, C. (2014). The impact of service characteristics on the switching costs–customer loyalty link. *Journal of Retailing*, *90*(2), 275–290. doi:10.1016/j.jretai.2014.04.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Boletín Oficial del Estado. (2018). *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*. Retrieved November 18, 2020, from https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673&p=20181206&tn=2

Boston Consulting Group & DEC. (2018). Informe BCG-DEC. In *Reinventarse: El gran reto de la Experiencia de Cliente. Estudio integral sobre la Experiencia de Cliente en España Evolución 2016-2018*. Retrieved November 28, 2020, from https://asociaciondec.org/informes-dec/informe-dec_bcg-reinventarse-el-gran-reto-de-la-experiencia-de-cliente/35680/

Capgemini. (2020). *Informe mundial del seguro 2020 de Capgemini*. Retrieved December 15, 2020, from https://www.capgemini.com/es-es/news/informe-mundial-del-seguro-2020-de-capgemini/

Carù, A., & Cova, B. (2013). *Consuming experiences: An introduction. In Consuming experience*. Routledge. doi:10.4324/9780203390498

CepymeNews. (2020). *14 estadísticas sobre la gestión de la experiencia del cliente.* Retrieved November 16, 2020, from https://cepymenews.es/gestion-experiencia-cliente/

Chang, K. C., Kuo, N. T., Hsu, C. L., & Cheng, Y. S. (2014). The impact of website quality and perceived trust on customer purchase intention in the hotel sector: Website brand and perceived value as moderators. *International Journal of Innovation, Management and Technology*, *5*(4), 255. doi:10.7763/IJIMT.2014.V5.523

Chen, C., & Phou, S. (2013). A closer look at destination: Image, personality, relationship and loyalty. *Tourism Management*, *36*, 269–278. doi:10.1016/j.tourman.2012.11.015

Cisco. (2020). *Cisco Annual Internet Report (2018–2023).* Retrieved November 21, 2020, from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

Cluetrain. (1999). *95 Theses*. Retrieved November 16, 2020, from https://www.cluetrain.com/; http://tremendo.com/cluetrain/

Cohen, H. (2013). *Los 4 momentos de la verdad en el marketing*. Retrieved December 13, 2020, from https://www.marketingdirecto.com/actualidad/checklists/los-4-momentos-de-la-verdad-en-el-marketing

Cohen, J. B., Pham, M.T., & Andrade, E.B. (2006): The nature and role of affect in consumer behavior. In Handbook of Consumer Psychology. Lawrence Erlbaum.

De Keyser, A., & Lariviere, B. (2014). How technical and functional service quality drive consumer happiness: Moderating influences of channel usage. *Journal of Service Management*, *25*(1), 30–48. doi:10.1108/JOSM-04-2013-0109

De Quevedo, E., De la Fuente, J. M., & Delgado, J. B. (2005). Reputación Corporativa y Creación de Valor. Marco Teórico de Una Relación Circular. *Investigaciones Europeas de Dirección y Economía de la Empresa*, *11*(2), 81–97.

Deloitte. (2018). *II Informe sobre la operativización de la Voz de Cliente.* Retrieved March 13, 2020, from https://www2.deloitte.com/us/en.html

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dhamija, P., & Bag, S. (2020). Role of artificial intelligence in operations environment: a review and bibliometric analysis. *Business Process Management Journal, 3*(1).

Duerig, T. W., Melton, K. N., & Stöckel, D. (2013). *Engineering aspects of shape memory alloys*. Butterworth-Heinemann.

Elsden, C., Manohar, A., Briggs, J., Harding, M., Speed, C., & Vines, J. (2018, April). Making sense of blockchain applications: A typology for HCI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). 10.1145/3173574.3174032

Findasense. (2020). Estudio COVID-19. *Escenarios y Horizontes de futuro post-cuarentena en el consumo y la producción.* Retrieved November 30, 2020, from https://es.insights.findasense.com/estudios/covid-19-paper-futuro-del-consumo-y-la-produccion-66508

Forrester. (2020). *Improving Customer Experience By One Point Can Drive More Than A Billion Dollars In Revenue*. Retrieved December 3, 2020, from https://go.forrester.com/blogs/improving-customer-experience-by-1-point-can-drive-more-than-a-billion-dollars-in-revenue-in-2019/

Forrester Customer Experience. (2013). *Forrester's 2013 Customer Experience Predictions.* Retrieved December 3, 2020, from https://go.forrester.com/blogs/13-01-04-forresters_2013_customer_experience_predictions/

Frijda, N. H. (1986). The emotions. Cambridge University Press.

García Haro, M. A. (2018*). La co-creación de valor a través de los social media: efecto en la imagen de Cuenca como destino turístico* (Doctoral thesis). Universidad de Castilla-La Mancha.

Gartner. (2019). *Creating a High-Impact Customer Experience Strategy*. Retrieved December 4, 2020, from https://emtemp.gcom.cloud/ngw/globalassets/en/marketing/documents/creating-a-high-impact-customer-experience-strategy-gartner-for-marketers-11-22-2019.pdf

Gartner. (2020). *Blockchain Technology: What's Ahead? Be ready for the next phase of the blockchain revolution.* Retrieved December 3, 2020, from https://www.gartner.com/en/information-technology/insights/blockchain

Gómez Borja, M.A. (2014). *El proceso de decisión en el consumidor*. UOC.

Guimerá, A. (2015). *Mapa de la empatía (Dave Gray).* Retrieved December 5, 2020, from https://www.marketing-esencial.com/2015/03/03/mapa-de-empatia/

Gutmann, G., & Konagaya, A. (2019). *Predictive Simulation: Using Regression and Artificial Neural Networks to Negate Latency in Networked Interactive Virtual Reality.* arXiv preprint arXiv:1910.04703.

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, *62*(5), 565–571. doi:10.1016/j.jbusres.2008.06.016

Hair, J., Hult, G., Ringle, C., & Sarstedt, M. (2017). *A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.). SAGE Los Angeles.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hendarman, A. F., & Tjakraatmadja, J. H. (2012). Relationship among soft skills, hard skills, and innovativeness of knowledge workers in the knowledge economy era. *Procedia: Social and Behavioral Sciences*, *52*, 35–44. doi:10.1016/j.sbspro.2012.09.439

iabSpain. (2018). *Estudio anual de eCommerce España 2018*. Retrieved December 6, 2020, from https://iabspain.es/wp-content/uploads/estudio-ecommerce-iab-2018_vcorta.pdf

Intersoft Consulting. (2020). *General Data Protection Regulation GDPR*. Retrieved December 6, 2020, from https://gdpr-info.eu/

ISO. (2020). *ISO/IEC 25010*. Retrieved December 18, 2020, from https://iso25000.com/index.php/normas-iso-25000/iso-25010

itUser. (2020). *La experiencia de cliente y el aumento de ingresos, una relación que confirman el 82% de las empresas.* Retrieved December 18, 2020, from https://www.ituser.es/estrategias-digitales/2020/06/la-experiencia-de-cliente-y-el-aumento-de-ingresos-una-relacion-que-confirman-el-82-de-las-empresas

Izard, C. E. (1977). *Human emotions*. Plenum Press.

Izo. (2019a). *5 Tendencias en Experiencia de Cliente en 2019*. Retrieved December 18, 2020, from https://izo.es/tendencias-experiencia-de-cliente-2019/

Izo. (2019b). *¿Cómo Conseguir Empleados Felices y Comprometidos?* https://izo.es/como-conseguir-empleados-felices-y-comprometidos/

Kleinginna, P.R., & Kleinginna, A.M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, *5*(4), 345–379.

Kohli, A. K., Jaworski, B. J., & Shabshab, N. (2019). Customer centricity: a multi-year journey. In *Handbook on Customer Centricity*. Edward Elgar Publishing.

Kotler, P., Armstrong, G., Cámara, D., & Cruz, I. (2004). Principios de marketing (10a edición). Madrid: Prentice Hall.

Lange, G. S., & Johnston, W. J. (2020). The value of business accelerators and incubators–an entrepreneur's perspective. *Journal of Business and Industrial Marketing*.

Lee, J., Singh, J., & Azamfar, M. (2019). *Industrial artificial intelligence.* arXiv preprint arXiv:1908.02150.

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*, 69–96.

Lin, C., & Lekhawipat, W. (2014). Factors affecting online repurchase intention. *Industrial Management & Data Systems*, *114*(4), 597–611. doi:10.1108/IMDS-10-2013-0432

Lindstrom, M. (2016). *Small data: las pequeñas pistas que nos advierten de las grandes tendencias*. Grupo Planeta Spain.

Lindstrom, M. (2019). *Brandwashed: O lado oculto do marketing controlamos o que compramos ou são as empresas que escolhem por nós?* Alta Books.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liuqu, Y., Fan, X., & Fu, P. L. (2015). From Customer Satisfaction to Customer Experience: Online Customer Satisfaction Practice in International E-commerce. In *International Conference on Cross-Cultural Design* (pp. 80-89). Springer. DOI: 10.1007/978-3-319-20934-0_8

Lorenzo Romero, C., Alarcón de Amo, M. C., & Gómez Borja, M. A. (2011). Adopción de redes sociales virtuales: Ampliación del modelo de aceptación tecnológica integrando confianza y riesgo percibido. *Cuadernos de Economía y Dirección de la Empresa*, *14*, 194–205.

Lourido Gómez, S., & Otero Neira, M.C. (2016). La dimensión afectiva de la experiencia de marca y el papel de las emociones en el comportamiento del consumidor. 26 *Congreso ACEDE Vigo, ponencia 3592ª*: [libro de actas].

Marketing Directo. (2019). *Las 4 tendencias en marketing digital que definirán el año 2020*. Retrieved December 18, 2020, from https://www.marketingdirecto.com/marketing-general/tendencias/las-4-tendencias-en-marketing-digital-que-definiran-el-ano-2020

Martín Lineros, E. (2020). *Webinar Alumni: La revolución del 5G*. Retrieved December 18, 2020, from https://www.youtube.com/watch?v=en8SzXmYmZU&feature=youtu.be

Martínez Sánchez, F. (2008). La emoción. In Motivación & emoción. Mc Graw Hill.

Mascarenhas, O. A., Kesavan, R., & Bernacchi, M. (2006). Lasting customer loyalty: A total customer experience approach. *Journal of Consumer Marketing*.

Medallia & DEC. (2020*). ¿Pueden ser emocionales las empresas financieras?* Retrieved December 18, 2020, from https://go.medallia.com/es-dec-webinar.html

Méndez Aparicio, M. D. (2019). *Determinantes y consecuencias de la Experiencia de Cliente y la Satisfacción en el ámbito digital: un análisis del uso de áreas privadas*© (Doctoral thesis). Universidad de Burgos, Burgos.

Méndez-Aparicio, M. D., Jiménez-Zarco, A., Izquierdo-Yusta, A., & Blazquez-Resino, J. J. (2020). Customer Experience and Satisfaction in Private Insurance Customer Web Areas. *Frontiers in Psychology*, *11*, 2591.

Méndez-Aparicio, M.D., Jiménez-Zarco, A.I., Izquierdo-Yusta, A., & Martinez-Ruiz. (2020). The psychological logic of customer loyalty: towards an integrated conceptual framework (satisfaction, trust and commitment). *Publicación en curso.*

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, *58*(Julio), 20–38.

Ning, L. Q., & Liu, J. L. (2008). Application of Stone-Geisser Test in Customer Satisfaction. *Tianjin Daxue Xuebao*, 3.

Ntwali, A., Kituyi, A., & Kengere, A. O. (2020). Claims Management and Financial Performance of Insurance Companies in Rwanda: A Case of SONARWA General Insurance Company Ltd. *Journal of Financial Risk Management*, *9*(03), 190.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Oh, H., Fiore, A. M., & Jeong, M. (2007). Measuring experience economy concepts: Tourism applications. *Journal of Travel Research*, *46*(2), 119–131.

Oliva, F. L., & Kotabe, M. (2019). Barriers, practices, methods and knowledge management tools in startups. *Journal of Knowledge Management*.

Palm, M. (2016). *Technologies of consumer labor: a history of self-service*. Routledge.

Pandian, A. P. (2019). Artificial intelligence application in smart warehousing environment for automated logistics. *Journal of Artificial Intelligence*, *1*(02), 63–72.

Pelegrín-Borondo, J., Arias-Oliva, M., & Olarte-Pascual, C. (2017). Emotions, price and quality expectations in hotel services. *Journal of Vacation Marketing*, *23*(4), 322–338.

Pelgrín Borondo, J. (2013) *Análisis comparativo de la estrategia de los grupos competitivos: e-clientes vs. clientes off-line* (Doctoral thesis). Universitat Rovira I Virgili, Tarragona.

Pérez-Rave, I. I., & Muñoz-Giraldo, L. (2014, July - September). Espacio literario relevante sobre la evaluación de la calidad del servicio: Países de realización de los estudios, métodos de análisis, índices de fiabilidad, hipótesis y desafíos. *Elsevier, Ingeniería. Investigación y Tecnología*, *15*(3). Advance online publication. doi:10.1016/S1405-7743(14)70356-7

Pilkington, M. (2016). Blockchain technology: principles and applications. In *Research handbook on digital transformations*. Edward Elgar Publishing.

Pine, J., & Gilmore, J. (1999). *The Experience Economy*. Harvard Business School Press.

Prebensen, N., Kim, H., & Uysal, M. (2015). Cocreation as Moderator between the Experience Value and Satisfaction Relationship. *Journal of Travel Research*, 1–12.

PuroMarketing. (2019). *En 2030 la Experiencia de Cliente dependerá de la Inteligencia Artificial*. Retrieved December 18, 2020, from https://www.puromarketing.com/12/32895/experiencia-cliente-dependera-inteligencia-artificial.html

PWC. (2017a). *La Inteligencia Artificial impulsará el PIB mundial un 14% en 2030 por sus efectos en la productividad y en el consume*. Retrieved December 18, 2020, from https://www.pwc.es/es/sala-prensa/notas-prensa/2017/la-inteligencia-artificial-impulsara-pib-mundial.html

PWC. (2017b). *Sizing the prize. What's the real value of AI for your business and how can you capitalise?* Retrieved December 8, 2020, from https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

PWC. (2020). *Experience is everything: Here's how to get it right.* Retrieved December 15, 2020, from https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/pwc-consumer-intelligence-series-customer-experience.pdf

Qualtrics. (2020). *What is closed-loop customer experience management?* Retrieved December 16, 2020, from https://www.qualtrics.com/experience-management/customer/closed-loop-cx/

Real Academia Española, R. A. E. (2019). *Diccionario de la lengua española* (22nd ed.). Retrieved December 17, 2020, from http://www.rae.es

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Reichheld, F. (2014). How the Net Promoter Score (NPS) Can Drive Growth. *JWI 518: Marketing in The Global Environment*, 73-100.

Rodríguez-Ardura, I., Maraver-Tarifa, G., Jiménez-Zarco, A. I., Martínez-Argüelles, M. J., & Ammetller, G. (2018). *Principios y estrategias de marketing* (Vol. 2). Edi UOC.

Russell, J. A. (2003). Core affect and psychological construction of emotion. *Psychological Review*, *110*(1), 145–172.

Sabio & DEC. (2020). *La omnicanalidad es responder a lo que el cliente te está pidiendo*. https://brilliantcx.sabiogroup.com/l/668023/2020-09-30/25ly6

Same, S., & Larimo, J. (2012, May). Marketing theory: Experience marketing and experiential marketing. In *7th International Scientific Conference. Business and Management* (pp. 10-11). Academic Press.

Sanny, L., Arina, A., Maulidya, R., & Pertiwi, R. (2020). Purchase intention on Indonesia male's skin care by social media marketing effect towards brand image and brand trust. *Management Science Letters*, *10*(10), 2139–2146.

Sarenet (2019). *2022: Más tráfico en Internet que el acumulado en tres décadas.* https://blog.sarenet.es/trafico-web/

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Sciences Information. Information Sur les Sciences Sociales*, *44*(4), 695–729.

Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management*, *15*(1-3), 53–67.

Shankar, V., Smith, A. K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, *20*(2), 153–175.

Shiau, W. L., Sarstedt, M., & Hair, J. F. (2019). Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Research*.

Sierra Diez, B., Froufe Torres, M., & Falces Delgado, C. (2010). El papel de las metas conscientes e inconscientes en la motivación del consumidor. *REME*, *13*(35-36), 193–210.

Skålén, P., Gummerus, J., Von Koskull, C., & Magnusson, P. R. (2015). Exploring value propositions and service innovation: A service-dominant logic study. *Journal of the Academy of Marketing Science*, *43*, 137–158. doi:10.100711747-013-0365-2

Spence, C. (2019). Neuroscience-inspired design: From academic neuromarketing to commercially relevant research. *Organizational Research Methods*, *22*(1), 275–298.

Swaid, S. I., & Wigand, R. T. (2012). The effect of perceived site-to-store service quality on perceived value and loyalty intentions in multichannel retailing. *International Journal of Management*, *29*(3), 301.

Taylor, S. J., & Bodgan, R. (1984). *Introducción. Ir hacia la gente. Introducción a los métodos cualitativos de investigación. La búsqueda de significados*. Paidós Ibérica.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Temkin, B. (2019). *The State of Customer ExperienceManagement, 2019. Benchmark of Customer Experience Activities, Competencies, and Maturity Levels*. Retrieved December 18, 2020, from https://www.qualtrics.com/docs/xmi/XMI_StateOfCustomerExperienceManagement-2019.pdf

Tokman, M., Davis, L. M., & Lemon, K. N. (2007). The WOW factor: Creating value through win-back offers to reacquire lost customers. *Journal of Retailing*, *83*(1), 47-64. doi:10.1016/j.jretai.2006.10.005

Torrent-Sellens, J. (2016). La economía del conocimiento y el conocimiento de la economía. *Oikonomics Revista de Economía. Empresa y Sociedad*, *5*, 26–32.

Torrent-Sellens, J. (2020). Industria 4.0 y resultados empresariales en España: un primer escaneado. *Oikonomics*. Retrieved December 18, 2020, from http://comein.uoc.edu/divulgacio/oikonomics/es/numero12/dossier/jtorrent.html

Triantafillidou, A., & Siomkos, G. (2014). Consumption experience outcomes: Satisfaction, nostalgia intensity, word of-mouth communication and behavioural intentions. *Journal of Consumer Marketing*, *31*(6/7), 526–540. https://doi.org/10.1108/JCM-05-2014-0982

Turletti, P. (2018). *El ROI de marketing y ventas, Cálculo y utilidad nuevo estandar de rendimiento*. ESIC.

Vargha, Z. (2018). Performing a strategy's world: How redesigning customers made relationship banking possible. *Long Range Planning*, *51*(3), 480–494.

Vargo, S. L., & Lusch, R. F. (2014). Inversions of service-dominant logic. *Marketing Theory*, *14*(3), 239–248.

Vázquez Brust. (2020). *Ciencia de Datos para Gente Sociable. Una introducción a la exploración, análisis y visualización de datos.* Retrieved December 18, 2020, from https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/index.html

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, *85*(1), 31–41.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, *28*(2), 118–144.

Walls, A. R., Okumus, F., Wang, Y., & Kwun, D. J. W. (2011). An epistemological view of consumer experience. *International Journal of Hospitality Management*, *30*(1), 10–21.

Wierenga, B. (2010). Marketing and artificial intelligence: Great opportunities, reluctant partners. In *Marketing intelligent systems using soft computing* (pp. 1–8). Springer.

Xiao, W., Zhao, H., Pan, H., Song, Y., Zheng, V. W., & Yang, Q. (2019, July). Beyond personalization: Social content recommendation for creator equality and consumer satisfaction. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 235-245). ACM.

Zarsky, T. Z. (2016). Incompatible: The GDPR in the age of big data. *Seton Hall Law Review*, *47*, 995.

Zhang, C. B., & Li, Y. N. (2019). How social media usage influences B2B customer loyalty: Roles of trust and purchase risk. *Journal of Business and Industrial Marketing*.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zomerdijk, L. G., & Voss, C. A. (2011). NSD processes and practices in experiential services. *Journal of Product Innovation Management*, *28*(1), 63–80.

ADDITIONAL READING

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2021). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, *114*, 106548. doi:10.1016/j.chb.2020.106548 PMID:32905175

Cachero-Martínez, S., & Vázquez-Casielles, R. (2021). Building consumer loyalty through e-shopping experiences: The mediating role of emotions. *Journal of Retailing and Consumer Services*, *60*, 102481. doi:10.1016/j.jretconser.2021.102481

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50. doi:10.100711747-020-00749-9

Kazak, A. N., Chetyrbok, P. V., & Oleinikov, N. N. (2020). Artificial intelligence in the tourism sphere. *IOP Conference Series. Earth and Environmental Science*, *421*(4), 042020. doi:10.1088/1755-1315/421/4/042020

Kumar, M. R., Venkatesh, J., & Rahman, A. M. Z. (2021). Data mining and machine learning in retail business: Developing efficiencies for better customer retention. *Journal of Ambient Intelligence and Humanized Computing*, 1–13. PMID:34025813

Mogaji, E., Soetan, T. O., & Kieu, T. A. (2020). *The implications of artificial intelligence on the digital marketing of financial services to vulnerable customers. Australasian Marketing Journal*.

Mustak, M., Salminen, J., Plé, L., & Wirtz, J. (2021). Artificial intelligence in marketing: Topic modeling, scientometric analysis, and research agenda. *Journal of Business Research*, *124*, 389–404. doi:10.1016/j.jbusres.2020.10.044

Purohit, S., & Jain, A. K. (2021). Leveraging IOT and AI for delivering tailor-made customer experiences in Indian petro retailing. *Indian Journal of Science and Technology*, *14*(7), 604–609. doi:10.17485/IJST/v14i7.3

Sánchez, C. A. C. (2021). Efectos de la inteligencia artificial en las estrategias de marketing: Revisión de literatura. *aDResearch. Revista Internacional de Investigación en Comunicación*, (24), 26–41.

KEY TERMS AND DEFINITIONS

Analytics: Intensive use of structured and unstructured customer data whose quantitative and qualitative statistical exploitation, predictive and explanatory models allow to improve the decision-making process of companies and the creation of competitive advantages and generation of value in organizations.

Customer Loyalty: Manifest intention of the client to maintain a continuous purchase or use relationship, choosing the brand over the competition.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Customer Satisfaction: Sentiment perceived by the client when their expectations are satisfied or exceeded by the company.

Customer Trust: The security perceived by the client that the company will not do anything that will harm their personal or economic status.

Globalization: Economic, technological, political, social, and cultural process on a world scale that connects the different countries of the world, uniting their social markets and causes deep social transformations.

Innovation: Modification of existing concepts in order to improve them or the creation of new concepts totally disruptive with the known reality. In marketing terms, it applies to new ways of satisfied customer needs.

Marketing 5.0: Offer consumer experiences in real time from a comprehensive knowledge of the customer and around the digitization of relationships based on the senses and human sensitivity.

Predictive Model of Consumer Behavior: Statistical model based on customer data that is capable of predicting their behavior in future consumer situations.

Small Data: Customer data from human understanding and of a qualitative nature, but that its volume and format make it accessible, informative and actionable for decision-making.

Voice of the Customer: Preferences, expectations, needs and opinions that the client expresses to the company in order to improve their interaction with the brand and the satisfaction of their needs.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 10
Using Intelligent Text Analysis of Online Reviews to Determine the Main Factors of Restaurant Value Propositions

Elizaveta Fainshtein
https://orcid.org/0000-0002-6503-7248
National Research University Higher School of Economics, Russia

Elena Serova
https://orcid.org/0000-0002-9510-3496
National Research University Higher School of Economics, Russia

ABSTRACT

This chapter discusses the sentiment classification of text messages containing customer reviews of an online restaurant service system using machine-learning methods, in particular text mining and multivariate text sentiment analysis. The study determines the structure of value proposition factors based on online restaurant reviews on TripAdvisor, collecting information on consumer preferences and the restaurant services in St. Petersburg (Russia) quality assessment and examines the influence of service format and reviews tonality on ratings restaurants factors. The service format context is proposed as the main attribute influencing the formation of the restaurant business value proposition and of relevance for online reviews. The results showed the key factors in the study of the sentiment were cuisine and dishes, reviews and ratings, and targeted search. MANOVA analysis represented that for special offers and features, reviews and ratings, factors and quantitative star ratings influenced the negative and positive sentiment of online reviews significantly.

DOI: 10.4018/978-1-7998-6985-6.ch010

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Nowadays, according to the increasing availability and popularity of e-commerce, social media, information portals, and media service platforms to gauge the opinions of restaurant consumers, online restaurant reviews are emerging as a new market phenomenon that plays an increasingly important role in value proposition decisions (Sheng et al., 2017; Ordanini & Pasini, 2008). Online reviews provide consumers with a wealth of information about service and product quality, which can reduce their uncertainty about ordering through restaurant systems. These reviews have a significant impact on restaurant service sales. With the ease of accessibility of mobile technology, online access through media service platforms to the experiences of consumers is accelerating and replacing more traditional forms of company brand assessment. The most popular online review platforms include sources such as Afisha Restaurants, Restoclub, TripAdvisor, Allcafe, and Restorating. The large volume of online restaurant reviews is also an indispensable resource for restaurants. Companies can analyze their customer experiences of service and food quality in more detail and in a timely manner, and understand potential growth points in relationship marketing and business modeling, whether there is a need to improve customer loyalty through online interaction, delivery systems, quality of service, or the menu, to determine the structure of significant factors in the value proposition.

The period of the pandemic served as a strong motivating aspect for the restructuring of commercial business and forced the transition to an online business format, and building and / or optimizing e-commerce systems and the delivery service, in a large number of commercial enterprises in the service sector, including the restaurant industry. With this in mind, it is important to understand what drives customers to come back or not to a restaurant, what motivates them to recommend a restaurant to their friends and family or not, what brand image is, and what factors create value for customers. In this regard, it is possible to track the growing popularity and importance of analytics for online restaurant service reviews, and the amount of relevant research using text mining and machine learning tools. A review of previous research related to online restaurant reviews published in academic research has shown an emerging reliance on Internet resources as sources of information for decision-making about restaurant products (e.g. Vásquez & Chik, 2015; Kaviya et al., 2017; Berezina et al., 2019), which heightens the need for more research on online restaurant reviews. It is important for a restaurant management system to use the information available from online customer reviews to better understand the needs of the target audience and improve the efficiency of the business. However, the online environment generates so much information that it can be difficult for managers to collect it all and analyze it manually. For this reason, this article uses text mining methods and systems that allow to extract meaningful patterns from large amounts of text information and demonstrate a possible way to process the data using machine learning (e.g. Hossain et al., 2017; Gogolev & Ozhegov, 2019; English & Fleischman, 2019; Ramos et al., 2020). The current study uses online customer testimonials, which are external evaluations of the restaurant business as part of e-commerce, a strong indicator of service quality. Detailed ratings provide more information about consumer preferences than individual overall ratings, and allow the timely modification of the restaurant's value proposition to increase the competitiveness of the service provided (Jannach et al., 2014). Multivariate analysis of the sentiment of the text allows to capture the subjectivity of online customer reviews in terms of semantic orientation associated with the constituents of the text.

In theory, identifying structured information from a free-form review text is difficult and time-consuming, so research on the main attributes of online restaurant reviews is rare. The field of research focused on understanding how factors specific to consumer inquiries in restaurants, as described in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

online customer reviews, reflecting their food experience, can influence a company's overall brand and brand ratings. Value proposition remains one of the topical areas where the amount of scientific research is only gaining momentum. Methodologically, most of the empirical research on online restaurant reviews typically uses only uninformative metadata, such as numerical ratings assigned by consumers, to avoid diving into the essential factors to improve restaurant business models. Previous research by the authors has shown that the semantic analysis of consumer requests provides richer and more extensive information for the restaurant industry and is a more informative indicator of consumer attitudes than the statistical processing of numerical ratings (Fainshtein & Serova, 2020). However, the format of the free text content of reviews in online reviews, in contrast to search queries, requires a separate study in the understanding of the authors.

Therefore, this study aims to close the knowledge gap by examining the underlying drivers of consumer demand for restaurant service, as described in online customer reviews, using machine learning. In particular, the research finds empirical answers to the following questions: What factors do consumers consider in online reviews when describing their experience of a restaurant? How do consumers' online attitudes about these factors affect overall restaurant ratings? This is achieved by comparing online restaurant reviews from satisfied customers who are willing to recommend the service to other consumers and dissatisfied reviews who discourage others from using a restaurant where they had a negative experience. The results of the study make it possible to understand which aspects of the food and service offered by restaurants generate positive and negative reviews in the online environment, which form the key factors of a competitive value proposition in the restaurant industry.

BACKGROUND: LITERATURE REVIEW

Customer Satisfaction With Online Restaurant Service and Behavioral Intentions

An online restaurant review includes describing and assessing the personal experience of a particular restaurant or purchasing meals or takeaways (Lane, 2013). Defining satisfied and dissatisfied customers is an important research issue among scholars of various disciplines, including management and marketing, which are actively used in the restaurant business (Titz et al., 2004; Parikh et al., 2014; Yang et al., 2017). The concept of measuring customer satisfaction and consumer behavior has been extensively studied by marketing researchers. The impact of social media on business transformation is recognized as important as it drives relationship marketing, brand sustainability and competitiveness. Repeat purchases and positive "electronic word of mouth" (EWOM) recommendations that encourage the use of a restaurant service directly depend on customer loyalty when writing online restaurant reviews. In other words, customer satisfaction reinforces a positive attitude towards a restaurant's online service brand and increases the likelihood of repeat purchases. On the other hand, dissatisfaction can lead to negative attitudes towards the brand and reduce the likelihood of repeat purchases.

At the moment, all online data about restaurants, their dishes and the service provided can be divided into two main groups: content created and posted directly by the producer of products and services (Producer Generated Content, PGC) and content created and posted directly by users (User Generated Content, UGC) independently of the restaurant (e.g. Salehi-Esfahani et al., 2016; Oliveira, & Casais, 2019; Cassar et al., 2020). The first group includes various marketing and advertising materials about

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the brand, products and services of the restaurant. The second group includes such types of social online content as forums, blogs, sites that systematize reviews, which may contain comments and consumer ratings, reviews, a collection of popular questions and answers (FAQ), user photo and video reports. The restaurant is able to independently influence the first type of content, choose the published information and form a personal brand, and then the second type directly depends on the influence of consumers, their assessments and the manifestation of loyalty to the restaurant, where the company is able to exert only an indirect influence.

One of the key approaches to answering the question of customer satisfaction and predicting a potential behavior strategy, which allows to form a high-quality value proposition, is measuring the service format (quality of service) and analyzing the key factors by which users make a choice when ordering online from a particular restaurant (Jensen & Hansen, 2007; Jin et al., 2012; Beuscart et al., 2016; Li et al., 2020). The service format is the level of service provision based on the perception of the client, which includes solving customer problems, satisfying the desires for a certain lifestyle, analyzing and controlling measures to create value for products, forming a competitive strategy, expressed in the initiation of emotions when selling goods and services for long-term customer support, stimulation of repeat purchases through a positive brand image. Elements of this definition have already been encountered in the papers of marketing and management researchers (Grönroos, 2011; Randhawa & Scerri, 2015). The perceived quality of service is part of a broader concept of customer satisfaction and behavioral intentions, which includes customer loyalty, which is key in shaping the value proposition.

Customers of online restaurant services use various elements to measure the quality of service they receive. The authors' research shows that in search queries, customer satisfaction is influenced by both the tangible and intangible aspects of service quality. In the previous study, after processing keyword data obtained in the Google KeyWord Planner system, keywords for the number of requests for restaurant service from restaurants in St. Petersburg (Russia) were broken down into seven main criteria for the most popular requests that can affect the attractiveness of the value proposition. Elements directly related to the service - geographic ((1) geolocation), psychographic ((2) reviews and ratings, including customer experience; (3) targeted search), behavioral ((4) promotions and discounts; (5) special offers and features), socio-demographic factors ((6) menu; (7) price). It has been argued that a denial of service or system failure that provokes less-than-full service can affect perceptions of service quality, satisfaction, and future behavioral intentions. Therefore, it is important for the restaurant business to recognize the factors that increase customer satisfaction and ensure their loyalty, on the basis of which a competitive value proposition can be formed.

The growth of e-commerce integration and the transition of the restaurant service system to online service increases the referral effect from regular customers. Dissatisfied online customer reviews can also be useful for restaurant management systems. First, comments like these can help restaurants by pointing out problem areas that may require careful attention and performance improvement (Harrison-Walker, 2019). Second, the impact of the service recovery paradox is also a significant factor when evaluating online reviews. The paradox of the restoration of services draws attention to the fact that the degree of satisfaction with customer service will be higher for those customers who, after specifying the defects in the service, received compensation for damage from the restaurant management than for those customers who immediately received the purchased services and products (Harrison-Walker, 2019; Żyminkowska, 2019; Mostert et al., 2012; Mejia et al., 2012). A review of scientific studies confirms that the service recovery paradox strategies increase customer loyalty as the company is attentive to feedback (Roggeveen et al., 2012; Liu et al., 2019). However, if the remarks are ignored or noticed but not addressed,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

this can lead to customer dissatisfaction with the service of the restaurant, and, consequently, to a low level of repeat purchases and negative EWOM (Yoo, 2020). A negative EWOM score causes significant damage to companies in the restaurant business, since most of the customers tend to leave and search for negative online reviews, considering negative comments to be more diagnostic, sincere and informative than positive or neutral ones.

In the decision-making process, reviews containing negative information are more important than positive information. A negative EWOM can deter potential customers from considering a particular product, service, or the restaurant brand itself, thereby damaging a company's reputation and financial strength. It can also very quickly go viral in online e-commerce systems and, spreading through all information channels, possibly reducing capital and brand image, reducing sales, and in extreme cases, completely shutting down the restaurant business, since the hospitality industry depends a lot on consumer loyalty.

Evaluating Customer Satisfaction With Online Restaurant Services in Web 3.0

The proliferation of electronic data collection systems, EWOM, and the development of e-commerce and the transition of most commercial brands to online sales during the pandemic, forever changed the status of customer opinion about food in online restaurant reviews. Based on Web 3.0 semantic technology, EWOM enables users to massively disseminate customer experiences of goods and services, as well as exchange and track content, which fosters a collective intelligence (Nayar, 2015). The semantic network in innovation for the exchange of data about EWOM services is manifested as the evolution of the value of consumer opinion, by taking into account the network interactions of users. Consumer-generated online reviews can be much more easily, systematically and widely disseminated than verbal, informal exchange of information between a small closed group of acquaintances or friends. More than ever, there is a trend that customers increasingly rely on online reviews to make decisions about which restaurant to order lunch from and which dish to choose (Ulker-Demirel, 2019). Therefore, it is not surprising that new paradigms of the technological possibilities of using the Semantic Web 3.0 allow a large number of customers to influence the value proposition of a restaurant. This makes it possible to make strategic managerial decisions, based on data analysis on the principle: the more users are involved, the more objective data can be collected and processed about the information material, products or services of the restaurant.

A survey of more than 5,500 online customers found that 59% of survey respondents considered consumer reviews to be more valuable than expert reviews (Piller, 1999; Kim et al., 2017). Online restaurant reviews have a broader impact on consumers, influencing demand, engagement marketing, and brand loyalty, both online and offline. In particular, online reviews focus managers on new opportunities for quality development and areas that need to be addressed through marketing innovation.

There are two forms of customer experience data analysis. The most popular form of online restaurant reviews is quantitative. A 1 to 5 star review system, is the most readily available to query. Moreover, such a form of input can be easily saved and processed by computer technology. In this way, a large dataset of customer testimonials can be collected over a long period of time, which can then be standardized and analyzed (Mustak et al., 2016).

Textual comments of online reviews are qualitative in nature. Descriptive data of customer experiences is a more complex and informative method requiring intelligent and multivariate analysis of the sentiment of the text. It reveals in detail what exactly had a positive or negative effect on the customer. Because online text reviews allow reviewers to substantiate the reasons and context, in which they

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

rated a product the way they did. This is one of the reasons why online text reviews usually complement numerical ratings. In other words, the contextual information contained in online text reviews can help potential customers make intelligent interpretations of the ratings from other users and therefore potentially make EWOM more useful and informative. Several websites, such as the online review service TripAdvisor, have been very successful in encouraging consumers to submit large numbers of text comments, reviews and ratings.

Online reviews of consumer experiences empower potential customers by enabling them to access more accurate and up-to-date information about the products and services of restaurants. In addition to customer interests, restaurant managers also benefit from working with information in online customer surveys, as they provide information on the strengths and weaknesses of the services provided, which makes text processing very useful in customer relationship marketing. Research suggests that customer feedback is increasingly important and influences consumer decision-making when choosing a restaurant (e.g. Yang et al., 2017; Kaviya et al., 2017; Chkoniya & Mateus, 2019; Li et al., 2020). Online customer reviews of products and services are becoming an increasingly important type of EWOM content. They are a valuable source of information to help customers make the right purchase decisions and help companies shape their value proposition in a service format more intelligently.

Factors That Affect Online Restaurant Review Rankings

Star ratings in online customer reviews play an important role in building consumer confidence in restaurant companies and are an important factor in business success. A number of research reviews have identified various factors that influence online rankings, such as post-transaction services (English & Fleischman, 2019), consumer experiences (Beuscart et al., 2016), relationship services (Berezina et al., 2019) and the use of reviewing media sites (Fan et al., 2017). However, there has been limited research on the factors that influence the ratings consumers give to restaurants through the use of text mining, and how consumer attitudes at the factor level affect their overall restaurant ratings.

Of particular note is the study by Zhang, Zhang, and Law (2014) showing the high impact of EWOM for restaurants, depending on four factors of restaurant service - food, physical environment, employee service, and cost. These factors greatly contribute to customer satisfaction in restaurants. Research has shown that one factor can replace another in the process of describing the value of a customer's use of a service. This research also confirms five criteria affect restaurant ratings in online reviews: food, service, atmosphere, price and context.

Therefore, the hypothesis for the study is that consumers' opinions of food, service, atmosphere, price and context have a significant impact on their overall positive or negative sentiment in restaurant rating text, as evidenced by the overall ratings in quantitative star values in online reviews of restaurant service.

METHODOLOGY

Data Collection

The data were taken from online restaurant reviews on TripAdvisor, which collects information on consumer preferences and the assessment of the quality of restaurants, St. Petersburg, Russia. The dataset consists of 201 customer reviews from 16 chain restaurants in St. Petersburg, Russia, including 242

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

restaurants. Seven main factors were included in the data analysis: factors based on value segments of consumer demand for restaurants in accordance with the classical segmentation structure: geographic (geolocation), psychographic (reviews and ratings, including impressions customers; targeted search), behavioral (promotions and discounts; specials and features), socio-demographic factors (menu; price) analyzed by the authors in a previous study (Fainshtein & Serova, 2020). Other categories in the data have been excluded to eliminate potential outliers.

Since the depth and granularity of an online review can affect the quality and quantity of information, and longer reviews include more data about products and services, how and where they were obtained, and in what context, the authors include in the analysis of each restaurant the 1 most recent review of at least 50 words. As a result, the dataset included 201 reviews from 242 restaurants in the 16 largest restaurant chains in St. Petersburg.

Predictive Analysis and Multivariate Analysis of the Text Sentiment

The growth of business information volumes means there is a need to analyze and automate different kinds of textual information.

Intelligent text analysis or text mining can be defined as automated extracting knowledge process from text data. Its peculiarity (in contrast to the analysis of other data) is that the initial information is not structured and formalized and it cannot be described by simple mathematical functions. The authors highlight three areas of the successful use of text mining applications and systems:

- supporting the decision making processes on the all levels of management;
- knowledge management;
- marketing research.

The main purpose of text mining is to extract information that is necessary for effective decision-making. Text mining allows not only the extraction of useful insights from unstructured data management projects, but also an expectation of higher ROI (return on investment) from them. For businesses, this means the ability to benefit from the use of big data, avoiding costly manual processing: to set aside irrelevant data and information and just get answers. With the growth of the volume of textual information to reduce the cost of its processing, there is a need for automation.

Intellectual capital is considered now as one of the main assets of an organization. It can generate up 50% of profit. As an organization grows, it accumulates a significant amount of intellectual assets. However, as a rule, their storage is not always well structured and standardized, different departments may use different tools for storing documents or even perform without storage. This makes it difficult or impossible to find information. The problem is especially real during mergers and acquisitions. To effectively use of intellectual capital, text-mining systems can:

- automatically collect and select information from various sources and compile it in a single format;
- provide an interface for searching documents by user-specified parameters;
- supplement documents with metadata (for example: source of documents, date of its creation, authors, etc.);
- pre-index and cluster documents;
- customize the levels of access to information based on security requirements.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Intelligent Text Analysis can be extremely useful in the sphere of marketing research. Text mining systems help to better understand the information field in which the company operates and how the customers (current and potential) relate to and cooperate with the company. Text mining systems can sort incoming customer's orders and provide more complete information about the customers and their needs at the output. The processing time of orders is reduced, saving money for the company.

To build a development strategy, companies need feedback and objective assessments of the use of their and competing products. Due to the large number of sources of information (scientific publications, product reviews, marketing research, conference proceedings, business news, etc.), automatic processing is also required here.

A very important task of text mining is sentiment analysis. The emotional attitude of the author needs to be evaluated. This is used, for example, to classify reviews of products or the assessment of the company itself.

Recent research in the field of marketing actively uses the technologies of intelligent and multivariate analysis of the sentiment of the text for analytics of marketing interactions and working with statistical data on consumer requests. Thanks to the methods of machine learning in information processing, the availability and efficiency of working with data from commercial organizations, including in restaurants, has increased.

The service format assumes that high-quality descriptive data of customer experiences in the text comments of online reviews are emotionally viewed as negative or positive, when certain factors in a certain tonal color encourage the client to leave comments about the strengths or weaknesses of the service. Therefore, in this study, the tonality of the text is understood as the attitude towards something described as a factor, negative or positive.

Intellectual and multidimensional analysis of the sentiment of the text (Kaviya et al., 2017; Hossain et al., 2017; Fan et al., 2017) is used to understand the attitude and opinion about the interaction with the company and customer's emotions about the product and service.

Sentiment analysis in the study was divided into two sequential stages: first, determining which text segments in the reviews contain dimensions and second, determining the polarity and strength of the tonality of each of these dimensions.

In the first phase of identifying attributes in online restaurant reviews, the study used factors from a previous study by the authors (Fainshtein & Serova, 2020) and developed a semi-supervised machine learning approach that included topic identification and mood classification. This approach identified seven attributes in online restaurant reviews that correlate strongly with positive and negative customer opinions, and helped determine both the overall scale of sentiment and the driving forces that restaurants need to consider when shaping their value proposition.

First, a set of keywords was proposed for each factor, for the initial loading of the classifier. The study used factors based on value segments of consumer demand for restaurants in St. Petersburg (Russia) in accordance with the classic segmentation structure. The authors build on their previous work (Fainshtein & Serova, 2020), where, using semantic analysis of keywords of consumer searches, the seven factors of the service format were identified, which potential customers pay attention to based on their requests for industry data. Each of the factor categories has fixed keyword attributes corresponding to the factor's topic.

Second, 201 online reviews from TripAdvisor's online review service were selected for analysis, each of which were the most recent text review containing at least 50 words, belonging to one of 242 restaurants in the 16 largest restaurant chains.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The seven factors were used to determine which of them were contained in the online survey. If the sentence contained any word that was included in one of the seven factor word lists, the sentence was considered to have the corresponding factor. A sentence could contain more than one word, and the words in a sentence could fall into more than one word list. Thus, the frequency of words that fell into the same list for each sentence was calculated. The list with the most words was identified as the main factor.

At the second stage of determining the polarity of emotional moods in the text (positive or negative sentiment), the study used sentiment lexicons in the QDA Miner Lite program and an extended list of 89 words compiled from the results of the CHAID analysis of consumer keywords (Fainshtein & Serova, 2020). The words were rated between most negative and most positive. The resulting QDA Miner Lite scores and sentiment scores are the sum of the scores for positive and negative words in a sentence. For multivariate analysis of the sentiment of the text, MANOVA was carried out using the SPSS 24.0 software package to check the influence of the seven factors and the quantitative assessments on the negative and positive sentiment of online reviews.

EMPIRICAL RESULTS

The sentiment analysis yielded negative and positive sentiments of online reviews (Table 1), and the seven indicators (Table 2) that measure the focus and depth of customer opinion on geolocation, reviews and ratings, targeted search, promotions and discounts, special offers, and features, menu, and prices of restaurants. To account for the different number of words in the reviews, the sentiment scores of the text (Table 2, Figure 1) were weighted in proportion to the number of sentences of each factor, based on the following calculations:

$$Sentiment_{ij} = Sentiment\ Score_{ij} \times \frac{Amount\ of\ words\ of\ the\ Factor_{ij}}{\Sigma\ Amount\ of\ words\ of\ the\ Factors_{ij}} (1)$$

where

i = 1 ... N online overview

Table 1. Sentiment of online reviews of restaurants in St. Petersburg, 2020

The sentiment of online reviews	Number of words	% of words	Tonality
Positive	17091	53.4	1.06
Negative	14895	46.6	0.93

j = geolocation, reviews and ratings, targeted search, promotions and discounts, special offers and features, menu, restaurant prices

This study took into account two control variables related to online customer reviews. Reviewing behavior is related to the emotional color of the review, which is posted on TripAdvisor, which shows the direction of the sentiment of online reviews (Table 1). Of great importance is the number of stars

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Significant factors in the value proposition of restaurants in online customer reviews in St. Petersburg, 2020

Factors	Number of words	% of words	Tonality
Geolocation	662	2.1	0.14
Reviews and ratings	8581	26.8	1.88
Targeted search	5817	18.2	1.27
Promotions and discounts	1873	5.9	0.41
Special offers and features	3920	12.2	0.86
Menu	9970	31.2	2.18
Price	1163	3.6	0.25

(from 1 to 5), which the client marked in the column for evaluating the restaurant's service on the site. The popularity of a restaurant chain is measured by the number of extreme positive and negative reviews received about individual restaurants within the chain. The sentiment targeting of online reviews was based on a binary approach: "1" for a restaurant with extreme positive reviews and "0" for negative reviews. The weighted ratings of the seven factors were used as explanatory variables, and the overall star ratings of each review and their sentiment direction were the dependent variable.

Next, were used the MANOVA model to test the relationship between the seven factors that measure the directionality of the sentiment of customers' opinions and the quantitative star ratings, accompanying text feedback on negative or positive sentiment of online reviews. Multilevel analysis allows to

Table 3. MANOVA analysis: testing the influence of seven factors and quantitative star ratings on negative and positive sentiment of online reviews

Multivariate results				
Factors	**df**	**Wilks's λ**	**F**	**Sig.**
Geolocation	25	0,220	0,048*	0,827
Reviews and ratings	103	0,219	8,530*	0,004
Targeted search	82	0,221	0,406*	0,525
Promotions and discounts	38	0,221	0,026*	0,871
Special conditions and features	69	0,220	14,931*	0,001
Menu	134	0,219	0,080*	0,777
Price	40	0,220	1,035*	0,310
* p‹0.05				

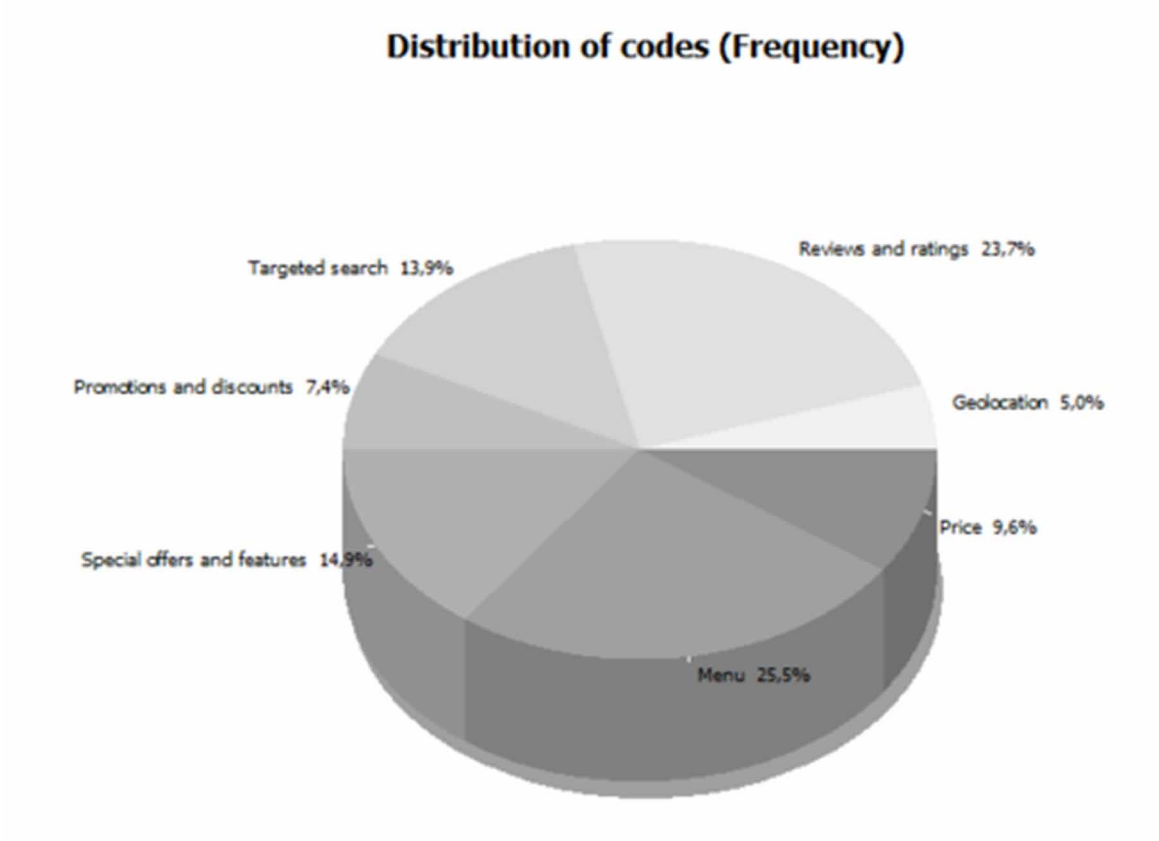
Note: the analysis uses the following values: df - degrees of freedom; Wilks's λ - Wilks lambda; F is the value of the F-criterion; Sig. - significance.

take into account in detail the hierarchical structure of the reviews and their settings, and to carry out a multidimensional analysis of the sentiment of the text.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The MANOVA analysis (Table 3) shows that for special offers and features, reviews and ratings, factors and quantitative star ratings influenced the negative and positive sentiment of online reviews. Special offers and features ($F = 14.931$, $p < 0.05$) and the criterion of reviews and ratings ($F = 8.530$, $p < 0.05$) have an impact on the value proposition within their group and confirm the hypothesis that consumers' opinions about food, service, atmosphere, price, and context have a significant impact on their overall

Figure 1. Distribution of words in online customer reviews of a St. Petersburg restaurant chain on TripAdvisor review platform, 2020



positive or negative sentiment in the restaurant review text, as evidenced by overall quantitative star ratings in online restaurant reviews. However, the criterion of geolocation, targeted search, promotions and discounts, menu, prices are not independent significant variables.

For online reviews from TripAdvisor, sentiment lexicons in the QDA Miner Lite program revealed the most influential factors in menu, reviews and ratings, and targeted restaurant search. For the subsequent MANOVA, the significant factors were special offers, reviews and ratings. This means that the majority of customers who reviewed on TripAdvisor visited the analyzed restaurants based on the reviews and ratings from other users, rather than responding to any other of the seven factors. In the first case, a study of the lexicons of sentiment shows that one group of respondents choose a restaurant based on the emotional component, where the review was most emotional, vivid and detailed (the number of words in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the review, which accounted for one of seven factors (Figure 1)), while another group made their choice based on a practical component, as shown by multivariate analysis of variance.

SOLUTIONS AND RECOMMENDATIONS

Theoretical Implications

This study used text mining to explain consumer behavior when writing online restaurant reviews and the factors that matter to them that restaurant managers could use in shaping the value proposition to increase competitiveness. The theory of text mining clearly substantiates the statistical significance of certain factors among online consumer reviews and shows that potential customers pay attention to positive and negative reviews when they evaluate information about the experiences of other consumers to avoid potential risks and reap the possible benefits of a restaurant visit. The theory also shows how to work with sentiment analysis tools to maximize restaurant visits by incorporating in-demand factors from online customer reviews into value propositions, thereby increasing demand for restaurant services.

This study demonstrates that a statistically significant positive relationship between positive or negative sentiment in restaurant rating text, seven service factors and quantitative star ratings in the restaurant business.

Previous scholars have conducted empirical research to test the impact of social media or traditional information technology ratings on restaurant performance by evaluating a focus on company performance. In terms of the key contribution, this study is the first to empirically identify key service format factors for value propositions and test the impact of online reviews that will help improve the future performance of restaurants.

In addition, with regard to considering the sentiment direction and ratings of online reviews on social media, this study collected online review information for each member of the largest restaurant chains in St. Petersburg (Russia) on TripAdvisor to improve reliability, as opposed to forming a variety of empirical findings from local websites or survey-based studies.

The theoretical contribution that this study makes is the result of the empirical finding that the feedback and rating factors, when customers share personal experiences from a restaurant have great power in explaining the performance metrics of a restaurant and the positive or negative direction of the text. This study is one of the few that looks at the rating of online reviews on social networks, identifying the service format factors for a value proposition in the restaurant business and analyzing which one has the most power influencing the purchase of service. Previous research has mainly focused on identifying social media attributes (Kaviya et al., 2017; Chkoniya & Mateus, 2019; English & Fleischman, 2019; Ramos et al., 2020). While social online review platforms have been around for a long time in restaurant marketing, restaurant marketers have used traditional information technology as the primary resource for business modeling and brand building for a long time. An analysis of online review ratings clearly shows that customer review ratings have strong predictive power in explaining the seven factors identified: geographic (geolocation), psychographic (reviews and ratings that include customer experiences; targeted searches), behavioral (promotions and discounts; specials and features), socio-demographic factors (menu; price). These results may offer a more informative value proposition resource for restaurant marketers and managers to predict the performance of their business model. A possible comparison of the relative magnitude of ratings across popular social media and online reseller reviews based on

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

restaurant commentary by experts clearly shows that TripAdvisor outperforms other social media and online restaurant review sites in predicting performance by providing richer, more recent information from users. This study also provides a comprehensive literature review of restaurant research, identifying the most popular and effective methods used for industry analytics.

Management Implications

The practical contributions of this study will help restaurant marketers understand the role of social media marketing and the potential application of online customer reviews. Restaurant marketers must consider both traditional customer satisfaction and the star rating of online reviews to maintain a competitive edge.

Social media can help restaurants by increasing the likelihood of table reservations by boosting EWOM promotion (Nayar, 2015). Restaurant managers can use social media platforms as an important brand building and engagement channel for developing customer relationships (Ulker-Demirel, 2019). Marketers should also track and use social media online review ratings as the dominant internal data to drive service and optimize value propositions. In addition, managers can expect to improve their overall financial performance by increasing their social media ratings. Getting higher online rankings leads to stronger brand awareness and more referrals from customers. Higher online review ratings, and working with the seven in-demand factors, can help reduce the cost of traditional promotion.

This study suggests that restaurant marketing strategists should continue to use online customer review tools to make management decisions. This will lead to better service, increased profitability, lower costs and increased customer loyalty.

Thanks to technological convergence, including the ability to collect a large amount of information about consumers, the analysis of data from online reviews is faster and more efficient. What's more, restaurants can now reach out to its customers in a more personalized way, which leads to an increase in the conversion of visits to sites with online reviews to offline sales. The restaurant must calculate its benefit from the existing business structure and value proposition by examining the potential costs and revenues, taking into account the factors important to the customer, by which he determines whether to visit a particular restaurant, and then make a decision to obtain the optimal combination of factors in his value proposition. This research helps marketers and restaurant managers better allocate marketing budgets and suggests a method to optimize the factors to include in the value proposition.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

While this study provides a comprehensive view of the positive or negative directionality of restaurant rating texts and star ratings for customer satisfaction, and identifies seven factors and their impact on restaurant value propositions, there are limitations. First, the use of predictive and multivariate analysis of the sentiment of the text may not be fully accurate. Second, future research can develop this theory and apply other analytical tools to study the relationship between customer satisfaction and their ratings on online social media reviews.

Further research could investigate whether the results of customer satisfaction reviews, social media ratings for the current period are more reliable than the results of the analysis of customer satisfaction in previous years. Researchers could also confirm the results of this study in hotel restaurants, in further research.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In addition, this study only included one of the most popular online review websites: TripAdvisor. Since there is the possibility that some new online intermediary or social media site could become a more significant influencer in the market, it is imperative that future research reflects newly emerging sites. Future research should expand the list of review sources to better understand the sentiment ratings of online reviews and their impact on restaurant value propositions.

CONCLUSION

This study examined the influence of consumers' opinions on food, service, atmosphere, price and context. The impact of seven factors typical for consumer inquiries in the field of restaurant services - geographic (geolocation), psychographic (reviews and ratings, including customer experience; targeted search), behavioral (promotions and discounts; special offers and features), socio-demographic factors (menu; price) on the positive or negative orientation of the sentiment of the restaurant rating text, and how quantitative star ratings in online restaurant service reviews are interconnected with them, which restaurant customers evaluate in their reviews.

According to the findings, ratings based on online restaurant reviews on TripAdvisor, which collects information on consumer preferences and the assessment of restaurants in St. Petersburg (Russia), are an informative source and allow to check the influence of seven factors and quantitative star ratings on the negative and positive tone of online reviews. The results confirm a direct link between the reviews and ratings factor, when customers describe their personal experience of interaction with the restaurant, the most significant thing to consider when forming a relevant, competitive value proposition for a restaurant business and focus on relationship marketing, customer focus and increasing the loyalty of potential clients.

This research can be useful for marketers, restaurateurs and their managers to quickly respond to consumer requests and modernize their value proposition based on which of the seven factors is their priority. The results of this study can directly increase the competitiveness of a restaurant and can improve marketers and managers' understanding of how to improve the quality of service for customers in general.

REFERENCES

Berezina, K., Ciftci, O., & Cobanoglu, C. (2019). Robots, artificial intelligence, and service automation in restaurants. In *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*. Emerald Publishing Limited. doi:10.1108/978-1-78756-687-320191010

Beuscart, J. S., Mellet, K., & Trespeuch, M. (2016). Reactivity without legitimacy? Online consumer reviews in the restaurant industry. *Journal of Cultural Economics*, *9*(5), 458–475. doi:10.1080/17530350.2016.1210534

Cassar, M. L., Caruana, A., & Konietzny, J. (2020). Wine and satisfaction with fine dining restaurants: An analysis of tourist experiences from user generated content on TripAdvisor. *Journal of Wine Research*, 1–16.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chkoniya, V., & Mateus, A. (2019). Digital Category Management: How Technology Can Enable the Supplier-Retailer Relationship. In Smart Marketing With the Internet of Things (pp. 139-163). IGI Global.

English, P., & Fleischman, D. (2019). Food for thought in restaurant reviews: Lifestyle journalism or an extension of marketing in UK and Australian newspapers. *Journalism Practice*, *13*(1), 90–104. doi: 10.1080/17512786.2017.1397530

Fainshtein, E., & Serova, E. (2020). Value Proposition of Network Companies Providing Restaurant Services in Russia: Analysis and Evaluation. In Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior (pp. 137-158). IGI Global.

Fan, Z. P., Che, Y. J., & Chen, Z. Y. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, *74*, 90–100. doi:10.1016/j.jbusres.2017.01.010

Gogolev, S., & Ozhegov, E. M. (2019). Comparison of Machine Learning Algorithms in Restaurant Revenue Prediction. In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 27-36). Springer.

Grönroos, C. (2011). A service perspective on business relationships: The value creation, interaction and marketing interface. *Industrial Marketing Management*, *40*(2), 240–247. doi:10.1016/j.indmarman.2010.06.036

Harrison-Walker, L. J. (2019). The effect of consumer emotions on outcome behaviors following service failure. *Journal of Services Marketing*, *33*(3), 285–302. doi:10.1108/JSM-04-2018-0124

Hossain, F. T., Hossain, M. I., & Nawshin, S. (2017). *Machine learning based class level prediction of restaurant reviews. In 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE.

Jannach, D., Zanker, M., & Fuchs, M. (2014). Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. *Information Technology & Tourism*, *14*(2), 119–149. doi:10.100740558-014-0010-z

Jensen, Ø., & Hansen, K. V. (2007). Consumer values among restaurant customers. *International Journal of Hospitality Management*, *26*(3), 603–622. doi:10.1016/j.ijhm.2006.05.004

Jin, N., Lee, S., & Huffman, L. (2012). Impact of restaurant experience on brand image and customer loyalty: Moderating role of dining motivation. *Journal of Travel & Tourism Marketing*, *29*(6), 532–551. doi:10.1080/10548408.2012.701552

Kaviya, K., Roshini, C., Vaidhehi, V., & Sweetlin, J. D. (2017). Sentiment analysis for restaurant rating. In *2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)* (pp. 140-145). IEEE.

Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction. *International Journal of Contemporary Hospitality Management*, *29*(2), 784–802. doi:10.1108/IJCHM-11-2015-0627

Lane, C. (2013). Taste makers in the "fine-dining" restaurant industry: The attribution of aesthetic and economic value by gastronomic guides. *Poetics*, *41*(4), 342–365. doi:10.1016/j.poetic.2013.05.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Li, H., Xie, K. L., & Zhang, Z. (2020). The effects of consumer experience and disconfirmation on the timing of online review: Field evidence from the restaurant business. *International Journal of Hospitality Management*, *84*, 102344. doi:10.1016/j.ijhm.2019.102344

Liu, H., Jayawardhena, C., Dibb, S., & Ranaweera, C. (2019). Examining the trade-off between compensation and promptness in eWOM-triggered service recovery: A restorative justice perspective. *Tourism Management*, *75*, 381–392. doi:10.1016/j.tourman.2019.05.008

Mejia, J., Mankad, S., & Gopal, A. (2020). Service Quality Using Text Mining: Measurement and Consequences. *Manufacturing & Service Operations Management*, 1–19. doi:10.1287/msom.2020.0883

Mostert, P., Petzer, D., & De Meyer, C. (2012). A theoretical and empirical investigation into service failure and service recovery in the restaurant Industry. In Service Science Research, Strategy and Innovation: Dynamic Knowledge Management Methods (pp. 86-99). IGI Global. doi:10.4018/978-1-4666-0077-5.ch005

Mustak, M., Jaakkola, E., Halinen, A., & Kaartemo, V. (2016). Customer participation management. *Journal of Service Management*, *27*(3), 250–275. doi:10.1108/JOSM-01-2015-0014

Nayar, R. (2015). Role of Web 3.0 in Service Innovation. In *The Handbook of Service Innovation* (pp. 253–280). Springer. doi:10.1007/978-1-4471-6590-3_13

Oliveira, B., & Casais, B. (2019). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*, *10*(1), 2–14. doi:10.1108/JHTT-11-2017-0130

Ordanini, A., & Pasini, P. (2008). Service co-production and value co-creation: The case for a service-oriented architecture (SOA). *European Management Journal*, *26*(5), 289–297. doi:10.1016/j.emj.2008.04.005

Parikh, A., Behnke, C., Vorvoreanu, M., Almanza, B., & Nelson, D. (2014). Motives for reading and articulating user-generated restaurant reviews on Yelp. com. *Journal of Hospitality and Tourism Technology*, *5*(2), 160–176. doi:10.1108/JHTT-04-2013-0011

Piller, C. (1999). Everyone is a critic in cyberspace. *Los Angeles Times, 3*(12), A1.

Ramos, K., Cuamea, O., Morgan, J., & Estrada, A. (2020). Social Networks' Factors Driving Consumer Restaurant Choice: *An Exploratory Analysis. In International Conference on Applied Human Factors and Ergonomics* (pp. 158-164). Springer.

Randhawa, K., & Scerri, M. (2015). Service innovation: A review of the literature. In *The handbook of service innovation* (pp. 27–51). Springer. doi:10.1007/978-1-4471-6590-3_2

Roggeveen, A. L., Tsiros, M., & Grewal, D. (2012). Understanding the co-creation effect: When does collaborating with customers provide a lift to service recovery? *Journal of the Academy of Marketing Science*, *40*(6), 771–790. doi:10.100711747-011-0274-1

Salehi-Esfahani, S., Ravichandran, S., Israeli, A., & Bolden, E. III. (2016). Investigating information adoption tendencies based on restaurants' user-generated content utilizing a modified information adoption model. *Journal of Hospitality Marketing & Management*, *25*(8), 925–953. doi:10.1080/19368623.2016.1171190

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*, 97–112. doi:10.1016/j.ijpe.2017.06.006

Titz, K., Lanza-Abbott, J. A., & Cruz, G. C. Y. (2004). The anatomy of restaurant reviews: An exploratory study. *International Journal of Hospitality & Tourism Administration*, *5*(1), 49–65. doi:10.1300/J149v05n01_03

Ulker-Demirel, E. (2019). The Features of New Communication Channels and Digital Marketing. In *Handbook of Research on Narrative Advertising* (pp. 302–313). IGI Global. doi:10.4018/978-1-5225-9790-2.ch026

Vásquez, C., & Chik, A. (2015). "I am not a foodie…": Culinary capital in online reviews of Michelin restaurants. *Food & Foodways*, *23*(4), 231–250. doi:10.1080/07409710.2015.1102483

Yang, S. B., Hlee, S., Lee, J., & Koo, C. (2017). An empirical examination of online restaurant reviews on Yelp. com. *International Journal of Contemporary Hospitality Management*, *29*(2), 817–839. doi:10.1108/IJCHM-11-2015-0643

Yoo, C. W. (2020). An Exploration of the Role of Service Recovery in Negative Electronic Word-of-Mouth Management. *Information Systems Frontiers*, *22*(3), 719–734. doi:10.100710796-018-9880-5

Zhang, Z., Zhang, Z., & Law, R. (2014). Relative importance and combined effects of attributes on customer satisfaction. *Service Industries Journal*, *34*(6), 550–566. doi:10.1080/02642069.2014.871537

Żyminkowska, K. (2019). Concepts of Customer Activism. In *Customer Engagement in Theory and Practice* (pp. 1–22). Palgrave Pivot. doi:10.1007/978-3-030-11677-4_1

KEY TERMS AND DEFINITIONS

EWOM (Electronic Word-of-Mouth): Any positive or negative review made by potential, current or former customers about a company's product that is available to many potential customers and is shared through social networks.

Multivariate Analysis of Variance (MANOVA): This is a type of multivariate analysis used for analyzing data that includes more than one dependent variable at a time.

Sentiment Analysis of the Text (Direction of the Review): Is the computer identification and clustering of the opinions expressed in a piece of text to determine what the author's attitude to a particular topic, product or service is (positive, negative, or neutral).

Service Format: The level of service provision based on the perception of the client, which includes solving the problems of the buyer of services, satisfying the desires for a certain lifestyle, analyzing, and controlling measures to create value for products, forming a competitive strategy, expressed in the emotions when selling goods and services for long-term customers, stimulating repeat purchases through a positive brand image.

Text Mining: An automated analytic tool for understanding and sorting unstructured text that makes it easier to manage the data it describes. Text analysis tools are often used to gain valuable insights in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

e-commerce and digital marketing, for example, when analyzing social media discussions, survey responses, and online reviews.

Text Mining Technologies: An artificial intelligence tool used by companies to transform raw data into useful information. By using software to find patterns in big data sets, businesses can learn more about their customers to develop better marketing strategies, increase sales, and reduce costs.

Value Proposition: A strategic tool for positioning a product or service related to the exceptional properties of the products, which, through its unique characteristics, helps the company differentiate itself from competitors in the market.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 11
Time Series Forecasting in Retail Sales Using LSTM and Prophet

Clony Junior
IEETA, University of Aveiro, Portugal

Pedro Gusmão
IEETA, University of Aveiro, Portugal

José Moreira
DETI, University of Aveiro, Portugal

Ana Maria M. Tome
DETI, University of Aveiro, Portugal

ABSTRACT

Data science highlights fields of study and research such as time series, which, although widely explored in the past, gain new perspectives in the context of this discipline. This chapter presents two approaches to time series forecasting, long short-term memory (LSTM), a special kind of recurrent neural network (RNN), and Prophet, an open-source library developed by Facebook for time series forecasting. With a focus on developing forecasting processes by data mining or machine learning experts, LSTM uses gating mechanisms to deal with long-term dependencies, reducing the short-term memory effect inherent to the traditional RNN. On the other hand, Prophet encapsulates statistical and computational complexity to allow broad use of time series forecasting, prioritizing the expert's business knowledge through exploration and experimentation. Both approaches were applied to a retail time series. This case study comprises daily and half-hourly forecasts, and the performance of both methods was measured using the standard metrics.

DOI: 10.4018/978-1-7998-6985-6.ch011

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The rise of data science as a discipline with application in very different fields drives many developments in data mining, machine learning (ML) or data visualization. One of the popular topics is time series (TS) forecasting, and the challenge is to find a model that fits a timely-ordered collection of past observations to estimate future values. Besides traditional statistical methods that have been used for a long time, new forecasting methods that include or are built entirely on ML techniques, such as decision trees, support vector machines and neural networks are being researched. However, TS forecasting is not an easy task. On one hand it requires in-depth knowledge of ML and statistical methods to use and optimize models of increasing complexity. On the other, domain knowledge is crucial for requirements elicitation, model selection, and results' interpretation. Typically, ML experts do not know the business in-depth, and domain experts do not understand ML. Therefore, how to combine the knowledge of professionals with these two profiles is an interesting question, which still needs answers.

The motivating case study in this work is the prediction of client visits in retail stores. TS forecasting in the field of retail mainly focuses on predicting products' demand using ARIMA, SVR, multilayer perceptron (MLP) and hidden Markov models (İşlek & Gunduz Oguducu, 2017; Maaß et al., 2014). However, the solutions used by current commercial systems to predict the number of visiting clients and sales are, for the most part, built on empirical methods. Thus, we present the methodologies of two recent and different TS forecasting approaches.

The first approach uses Long Short-Term Memory (LSTM), a special kind of Recurrent Neural Network (RNN) that can deal with long sequential data. Some ANN features are challenging to achieve using other TS forecasting approaches because ANNs are self-adaptive, support nonlinear models and deal with multivariate problems. LSTM can learn which temporal dependencies in the data should be kept and which should be forgotten while enforcing mechanisms to reduce short-term memory. The reason for choosing the LSTM, however, comes from its forecasting potential as there are studies that demonstrate clear superiority of LSTMs over classical models such as ARIMA, a classical model that is very much established in the TS forecasting scene (Kolarik & Rudorfer, 1994; Siami-Namini et al., 2019).

The second approach uses Prophet, an open-source library developed by Facebook for time series forecasting. Prophet uses a generalized additive model (Hastie & Tibshirani, 1987) to represent trend, seasonality and holidays. While LSTM forecasting models are solely developed by experts in data mining or ML, Prophet aims to use general-purpose methods to keep domain experts in the loop. This includes understanding the data and the problem, choosing and tuning the forecasting model, and validating the model using adequate metrics. These issues are explained and discussed by example in this chapter.

The target audience is practitioners and academics interested in understanding and implementing TS forecasting solutions. The case study consists in predicting the number of client visits in retail stores, but the contents and issues presented in this chapter are also of interest in other use cases. The focus is on presenting the methodology, issues, and possible solutions using two different approaches in TS forecasting and not on optimizing or comparing the forecasting models.

The chapter is organized in the following order: (1) background in time series forecasting using neural networks and deep learning (LSTM), and off-the-shelf tools for business scenarios (Prophet) ; (2) the data and case study; (3) methodology and the implementation of the case study using LSTM and Prophet; (4) illustrative results and discussion of the two implementation strategies; (5) conclusions and guidelines for future research.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

TS forecasting is an important task that helps organizations with capacity planning. For instance, in the retail business, sales forecasting is behind all strategic and planning decisions. An accurate sales forecast helps with efficient inventory and logistics management.

A TS is a sequence of values measured at successive time instants. The time elapsed between the measurements is usually equally spaced, e.g., minutes, hours, days or months. TS can have three types of components:

- **Trend.** A trend exists when there is a general tendency of a time series to increase, decrease or stagnate over time. For instance, the increase or decrease of the unemployment rate, mortality rate or the number of houses in a city can hold for a considerable time.
- **Seasonal.** A seasonal pattern occurs when the data is influenced by factors every year, month, day of the week, etc. For instance, ice-cream sales increase in summer, woolen clothes sales increase in winter and the number of visitors to a museum increases during weekends.
- **Irregular.** This component represents random variations of a time series caused by unpredictable factors or incidences, such as strikes, floods, etc. There is no statistical method for measuring random fluctuations.

There are two preconditions for using TS forecasting methods: 1) there must be numerical data about the past, and 2) it should be reasonable to assume that past patterns will continue in the future. Many forecasting methods exist, some specializing in specific problems, and the most appropriate one for a given use case depends mainly on the data. This section presents two methods that will be used later to implement the same problem.

Long Short-Term Memory Neural Networks

The computational models of primitive ANNs were inspired by the brain's neural activity (Rosenblatt, 1957). These pioneer proposals are simple linear models since efficient algorithms to optimize multilayer networks' connections were missing. However, with the application of the backpropagation algorithm proposed in ANNs (Werbos, 1982) these models have been accomplishing great achievements in different fields, like medicine, economy, and so on (Abiodun et al., 2018). During the last decade, with the availability of large datasets and computing power, these models have evolved to efficient and versatile computational models. Modern architectures like GoogleNet (Szegedy et al., 2015) are composed of repeatable units. These units are made up of several blocks that can have different functionalities. The Long Short-Term Memory (LSTM) is an example of such architecture.

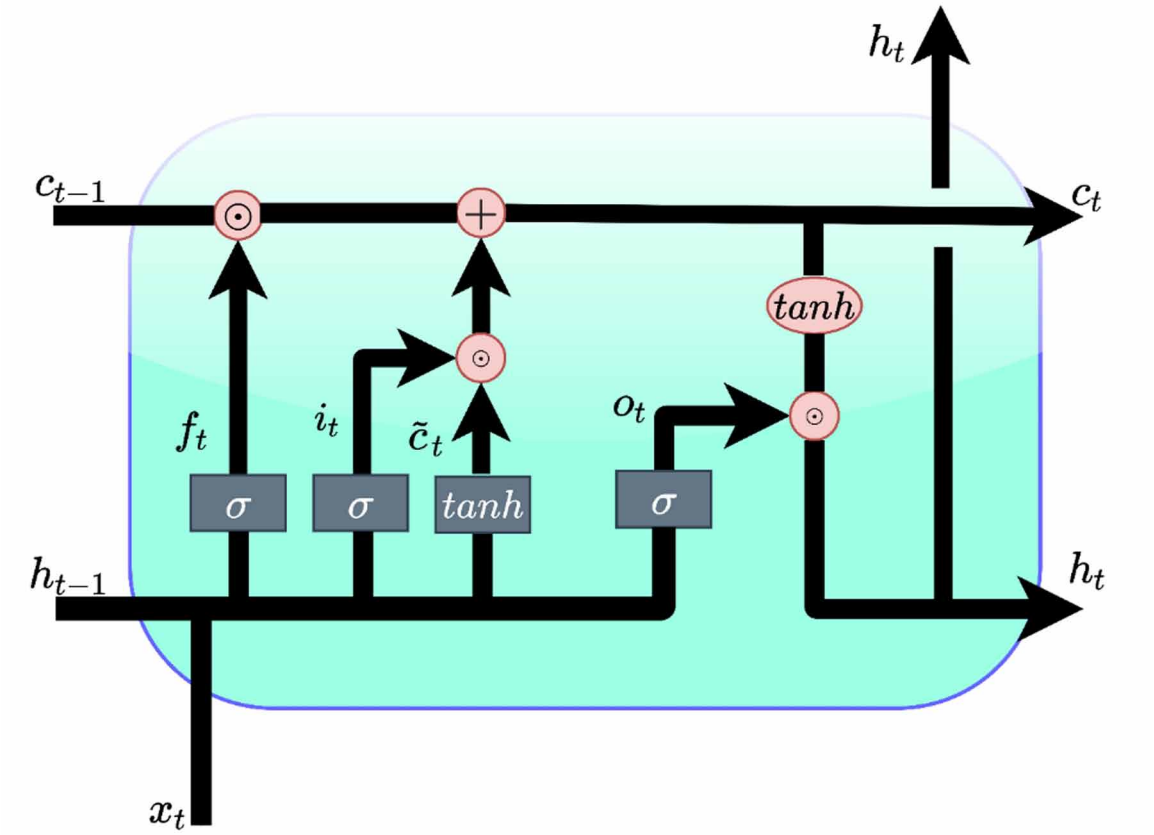
LSTM is a special variation of the Recurrent Neural Network (RNN). RNNs belong to a distinctive type of ANNs designed to model problems with a sequential or time-related aspect. RNNs can "memorize" by preserving an internal state using feedback loops. These loops occur in the network's neural units; thus, they are called recurrent neural units. In contrast to standard ANNs, a recurrent neuron has a connection to itself, making it possible to use past outputs.

When RNNs were first proposed (Hopfield, 1982), their popularity was short-lived because of the vanishing gradient problem occurring during ANN training (Bengio et al., 1994 and Gers et al., 2000), which is even more frequent with RNNs. LSTM was proposed to overcome this problem (Hochreiter &

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schmidhuber, 1997). The overall control flow of an LSTM is similar to an RNN, differing only in the repeating module, i.e., the recurrent unit's internal processes.

Figure 1. Recurrent processes in an LSTM unit. The grey rectangles refer to the recurrent activations



An LSTM unit has three main concepts (*Figure 1*):

- The cell state $\mathbf{c}_t$, consists of an encoded version of the information gathered from all the previously processed timesteps. As a result, this term is also called the long-term state;
- The block output $\boldsymbol{h}_t$, is similar to the cell state, but it only holds information from the previous timestep. Thus, it is also called the short-term state;
- The gating mechanisms are called input gate ($\boldsymbol{i}_t$), output gate ($\boldsymbol{o}_t$) and forget gate ($\boldsymbol{f}_t$). The input gate determines what information from the candidate memory $\left(\tilde{\boldsymbol{c}}_t\right)$ should be appended to the cell state. The output gate controls which values from the cell state $\left(\mathbf{c}_t\right)$ should be outputted within $\boldsymbol{h}_t$. The forget gate $\left(\boldsymbol{f}_t\right)$ determines what information should be removed or "forgotten" from the cell state $\left(\mathbf{c}_t\right)$.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this study, the input is a vector $\boldsymbol{x}_t$,

$$\boldsymbol{x}_t = (x(j-p+1), x(j-p+2), \ldots, x(j))^T \tag{1}$$

formed by embedding p samples using a sliding window on a TS $x(j)$ with $j \in p-1, \ldots, J-H-1$, such that J is the number of training records of $x(t)$ and H is the forecasting horizon. The LSTM unit can be expressed following an intuitive order of operations, where all gates and the candidate memory share the same format. They all deal with a combination of the input $\boldsymbol{x}_t$, the previous block output $\boldsymbol{h}_{t-1}$, and $\boldsymbol{W}$ and $\boldsymbol{R}$ refer to the matrices of weights of the connections. The forget gate can be formally written as

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}^{(f)}\boldsymbol{x}_t + \boldsymbol{R}^{(f)}\boldsymbol{h}_{t-1} + \boldsymbol{b}^{(f)}\right) \tag{2}$$

The symbol A refers to a specific activation function, e,g, a sigmoid, that transforms all values into $]0,1[$. If the values are close to zero, the data is to be removed, and if they are close to one, then it is to be kept or added.

In the next step, the unit decides what information is to be appended to the cell state $(\boldsymbol{c}_t)$. First, the input gate $(\boldsymbol{i}_t)$ decides which values should be updated, and second the new candidate values $(\tilde{\boldsymbol{c}}_t)$ that should be appended to the cell state are created. This is formally written as

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}^{(i)}\boldsymbol{x}_t + \boldsymbol{R}^{(i)}\boldsymbol{h}_{t-1} + \boldsymbol{b}^{(i)}\right) \tag{3}$$

$$\tilde{\boldsymbol{c}}_t = tanh\left(\boldsymbol{W}^{(\tilde{c})}\boldsymbol{x}_t + \boldsymbol{R}^{(\tilde{c})}\boldsymbol{h}_{t-1} + \boldsymbol{b}^{(\tilde{c})}\right) \tag{4}$$

where *tanh* is an activation function (hyperbolic tangent) and transforms all values into $]-1,1[$. Next, the cell state is updated from $\boldsymbol{c}_{(t-1)}$ to $\boldsymbol{c}_{(t)}$, where the $\odot$ operator represents the element-wise product.

$$\boldsymbol{c}_t = tanh(\boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{c}}_t \tag{5}$$

The last steps assemble the unit's output. First, the output gate decides through the sigmoid function which values of the updated cell state should be carried on in the block output $\boldsymbol{h}_t$. Second, the element-wise product of the output gate and the updated cell state resulting vectors are calculated. This is written as,

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}^{(o)}\boldsymbol{x}_t + \boldsymbol{R}^{(o)}\boldsymbol{h}_{t-1} + \boldsymbol{b}^{(o)}\right) \tag{6}$$

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\widehat{\boldsymbol{y}}_t = \boldsymbol{h}_t = \boldsymbol{o}_t \odot \boldsymbol{c}_t \tag{7}$$

where, in this context, the block output $\boldsymbol{h}_t$ becomes the model's output, i.e., the forecast $\widehat{\boldsymbol{y}}_t$ at the last processing step.

The training of standard ANNs is done using the backpropagation algorithm. However, for a RNN such as the LSTM an extension is necessary. For that goal, the backpropagation through time (BPTT) was proposed in (Gers et al., 2000). Fundamentally, it is processed as the backpropagation algorithm but while considering every timestep t.

Prophet

TS data in business problems tend to hold several common features and follow specific patterns, making it practical to develop specialized software packages for the sector. This is the case of Prophet, a free software package released by Facebook to study time series data. Besides TS forecasting, it also includes visualization tools and cross-validation facilities to measure forecasting errors using historical data.

Prophet uses a generalized additive model (Hastie & Tibshirani, 1987) to fit the time series. Therefore, the model is described by additive components representing trend ($g(t)$), seasonality ($s(t)$) and holidays ($h(t)$):

$$x(t) = g(t) + s(t) + h(t) + \epsilon(t) \tag{8}$$

where $\epsilon(t)$ denotes the residuals or any properties that are not fitted in the model.

Trend Model

Prophet implements two trend models: a linear model and a nonlinear saturating growth model. Both models use a piecewise strategy by incorporating changepoints, i.e., points where the growth rate can change. Changepoints are defined as $(s_j, \delta_j), j = 1, 2, \cdots S$, where δ_j is the change in the growth rate at s_j. The vector $'$ denotes all growth rate changes of the time series. Thus, the growth rate at any time t is the sum of all growth rate changes before t, calculated as follows. Let $\boldsymbol{a}(t)$ denote a vector with binary entries $\{0,1\}$ such that $a_j(t) = 1$, if $t \geq s_j$, otherwise $a_j(t) = 0$. The growth rate at time t is $k + \boldsymbol{a}(t)^T{}'$, where k is the initial logistic growth rate.

Nonlinear saturating growth with changepoints model

In this case, trend is a logistic saturating growth function and so it has the following template:

$$f(t) = \frac{L}{1 + e^{-k(t-m)}} \tag{9}$$

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

where L is the maximum in the curve, k is the logistic growth rate and m is the value of t at the sigmoid's midpoint. The nonlinear saturating piecewise growth model is defined as:

$$g(t) = \frac{c(t)}{1 + e^{-\left(k + \boldsymbol{a}(t)^T \boldsymbol{\delta}\right)\left(t - \left(m + \boldsymbol{a}(t)^T \boldsymbol{\gamma}\right)\right)}} \tag{10}$$

where $\boldsymbol{c}(t)$ is the saturation or carrying capacity. It is a business-related parameter, e.g., market size, population or GDP, which can change over time. The vector $\boldsymbol{\gamma}$ is used to adjust the values of m (see Equation (11)) to achieve the continuity of the function at the changepoints as follows:

$$\gamma_j = \left(s_j - m - \sum_{l<j} \gamma_l\right)\left(1 - \frac{k + \sum_{l<j} \delta_l}{k + \sum_{l \le j} \delta_l}\right) \tag{11}$$

Linear with changepoints model

This model is used when the time series does not incorporate any saturating growth and the trend is defined as

$$g(t) = \left(k + \boldsymbol{a}(t)^T \boldsymbol{\delta}\right)t + \left(m + \boldsymbol{a}(t)^{\tau} \boldsymbol{\gamma}\right) \tag{12}$$

where the entries of the adjustment vector are $\gamma_j = -s_j \delta_j$.

Changepoint selection

The changepoints can be assigned by a business expert, e.g., by entering the starting date of promotions, or calculated automatically from an initial grid of equally space changepoints defined over the whole history T of the time series. The number of changepoint candidates is 25 by default. The selection of changepoints is achieved using a prior $\delta_j \sim Laplace(0, \tau)$ in Equation (12). This imposes a tradeoff between minimizing the Mean Squared Error (MSE) between the model and the data and setting as many δ_j to zero as possible. When τ tends to zero, the trend component becomes a non-piecewise model ($j = 1$).

Seasonality

Seasonality components are modeled using Fourier Series. The amplitudes of the sines and cosines are the parameters to be adjusted and the periods (hourly, daily, weekly, monthly, etc.) are adjusted by the users. Fitting the seasonality is defined as

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$s(t) = \sum_{k=1}^{K} \left(a_k \cos\left(\frac{2\pi kt}{P} \right) + b_k \sin\left(\frac{2\pi kt}{P} \right) \right) = \boldsymbol{X}(t) \, ²$$ (13)

In this formulation, the sum is replaced by an algebraic formula where the sinusoidal function is written as entries of a row matrix $\boldsymbol{X}(t)$ and the column vector $² = (a_1, a_2, \cdots a_K, b_1, b_2, \cdots, b_K)^T$. P is a user-defined parameter for seasonality. For instance, $P = 7$ means weekly, $P = 365.25$ means yearly, etc. The number of terms (K) should not be too large to avoid overfitting. The documentation of Prophet recommends using $N = 3$ for weekly and $N = 10$ for yearly seasonality.

Holidays and Events

Prophet allows one to configure a parameter with a list of past and future events that contemplate known holidays and local events affecting the business. The use of the holiday list in the model was designed to make the process as simple as possible. Let D_i be the set of past and future holidays. An indicator function is added that represents whether time t is during a holiday i, and it is assigned each holiday a parameter $\varkappa_i$ which is the corresponding change in the forecast. This is done in a similar way to seasonality, generating a matrix of regressors (Taylor & Letham, 2018).

$$\boldsymbol{Z}(t) = \left[1(t \in D_1), \cdots, (t \in D_L) \right]$$ (14)

and taking,

$$h(t) = \boldsymbol{Z}(t) \varkappa$$ (15)

Prophet also offers a means of including the effects for the days around a specific holiday. There are additional parameters to delimit the days surrounding a holiday and basically, each of these days are considered holidays themselves.

The trend uncertainty is estimated by assuming that the future will see the same average frequency and magnitude of rate changes seen in history. As projected forward, it will produce wide uncertainty intervals (Taylor & Letham, 2018).

Metrics

The evaluation of a quantitative forecast is usually treated as a measure of how accurate or how well a model fits or reproduces the data. For that, various statistical error measures that give different insights were created. Before detailing some of the existing metrics, a definition of error is important. Formally, the error is

$$e(t) = y(t) - \hat{y}(t)$$ (16)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

where $y(t)$ is the observed value and $\hat{y}(t)$ the value forecasted both for the same time t. It is worth noting that this definition is allocated in a single-step forecast scenario, meaning that the forecasted value $\hat{y}(t)$ was built on observations before $y(t)$. Thus, $e(t)$ is a one-step forecast error.

However, forecasting methods usually produce forecasts with more than one step, i.e., feature a forecasting horizon $H > 1$. The following standard statistical accuracy metrics can be used:

$$\begin{aligned}
\text{Mean Error} &= \frac{1}{H}\sum_{t=1}^{H} e(t) \\
\text{Mean Absolute Error} &= \frac{1}{H}\sum_{t=1}^{H} |e(t)| \\
\text{Mean Square Error} &= \frac{1}{H}\sum_{t=1}^{H} e(t)^2 \\
\text{Root Mean Square Error} &= \sqrt{\frac{1}{H}\sum_{t=1}^{H} e(t)^2}
\end{aligned} \tag{17}$$

The mean forecast error is simply the average of the errors for all H timesteps. Usually, this metric is not very useful since it does not provide a clear error description when there are negative and positive values (they cancel each other out). However, this metric can suggest whether the model is consistently over forecasting or under forecasting. That is why the mean forecast error is also referred to as forecast bias. The other metrics circumvent this issue by molding the error to a positive value. Before averaging, the MAE metric turns the error into a positive value by taking its absolute value, while both MSE and RMSE metrics achieve that through squaring.

A problem that is common to these metrics is that they are not unit-free, i.e., they depend on the scale of the given data. Therefore, alternative metrics are required to render possible the comparison of forecast performances across TSs with different units. With that goal, metrics that provide relative errors are the usual approach. The relative error is given as,

$$p(t) = \left(\frac{y(t) - \hat{y}(t)}{y(t)} \right) \times 100 \tag{18}$$

Due to their intuitive interpretations and high popularity, MAPE (Equation (19)) and the absolute error metrics MAE and RMSE are the chosen evaluation functions in this work.

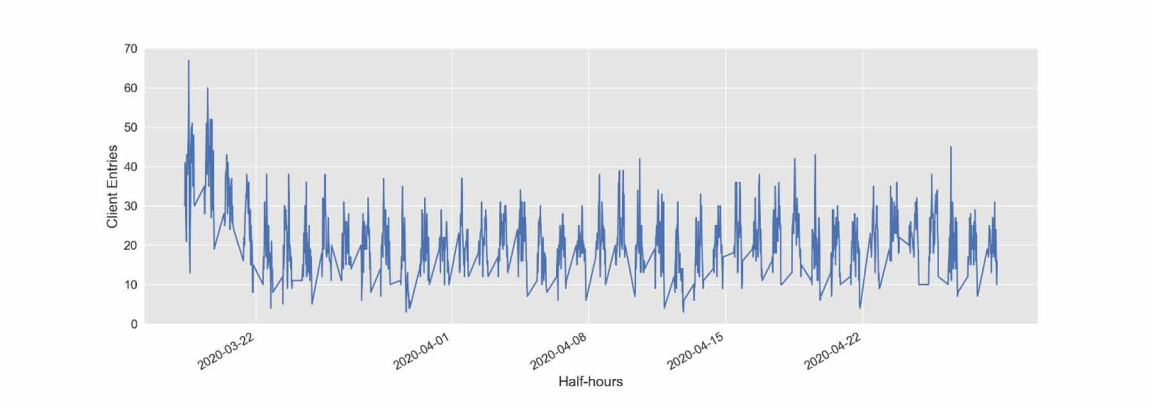
$$\text{Mean Absolute Percentage Error} = \frac{1}{H}\sum_{t=1}^{H} |p(t)|$$

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CASE STUDY

The case study consists in predicting the number of client entries in a retail store, by day and every 30 minutes, within a horizon of 28 days. The TS consists of records with the number of clients who entered the store between 2020-03-18 and 2020-05-26, every 30 minutes (Figure 2).

Figure 2. Number of client entries every half-hour



During the first two days (March 18 and 19, 2020), the number of clients is greater than during the remaining days in the time series. This period coincides with the beginning of the lockdown due to COVID-19. The store is open 7 days a week and the small gaps in the TS represent the time intervals when the store is closed. A closer observation of the data showed that the opening schedule was not the same for all days of the week and that even for the same day of the week, there were changes during that time. To make the formulation of the problem easier, some samples were deleted to have the same schedule for all stores. In the new dataset, all samples are from 08:30 to 20:30. Figure 3 displays the average number of client entries by half-hour.

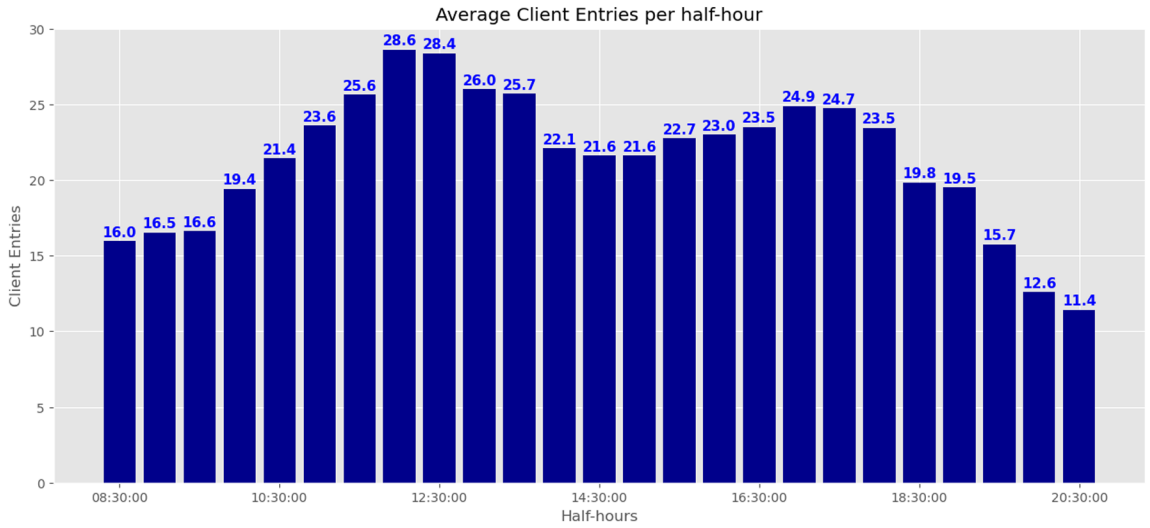
Two datasets were used in this study, the original and a new one derived from this one, with the data aggregated by day. In both cases, the TS is split into train and test datasets. The training dataset has 1050 observations (42 days) and the test dataset has 700 observations (28 days). The first two days of the time series, where the number of clients differs from the usual due to the lockdown, are in the training set, influencing the forecasting results. Figure 4 shows the average number of clients per day of the week. The days that show less clients are at Mondays and Tuesdays, i.e., the beginning of the week. Thursday was the weekday where more clients entered the store.

EXPERIMENTAL SETUP

This section presents the implementation of the case study using the two approaches studied in this chapter. The first, LSTM, required in-depth analysis to evaluate different configurations to fit the TS model. The second, Prophet, is already prepared to deal with this kind of problem and the main task is to obtain the properties of the TS using special-purpose parameters.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Average number of client entries in each half-hour

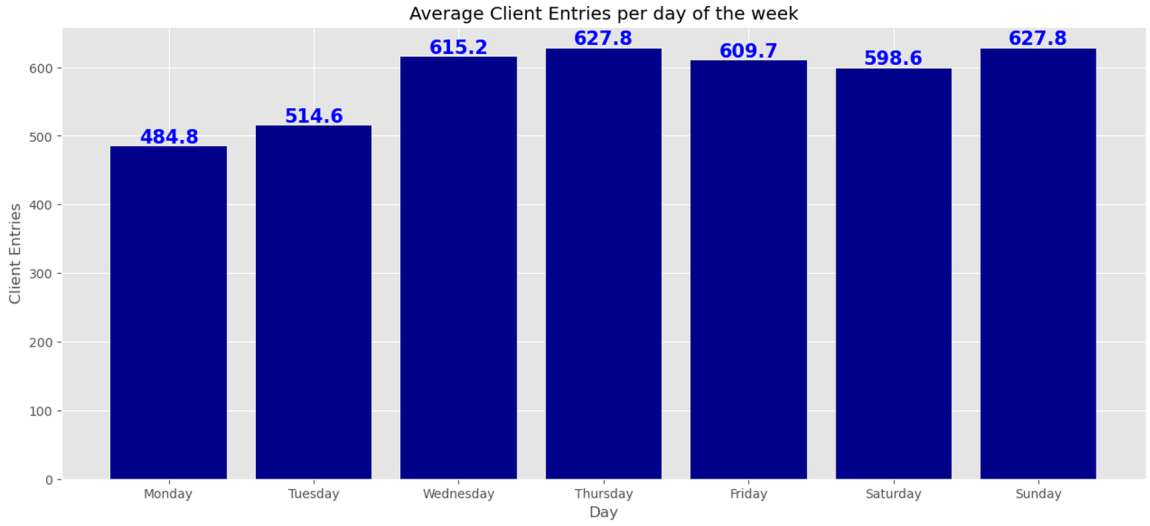


LSTM

Technology and Libraries

The neural model and related operations were developed using Python 3.7.9. The following libraries were used:

Figure 4. Average number of client entries in each day of the week



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- *Pandas*[1] was used to extract the data and to split the data into training and test sets.
- *Numpy*[2] provides efficient and optimized functions to deal with multidimensional arrays. This is important as these arrays are required for the network's operations.
- *TensorFlow*[3], developed by the *Google Brain Team*, is an open-source library for distributed numerical and auto-differentiable computations. These computations are the primary support for many developed deep learning algorithms. It was initially implemented in C++, but a native Python API is also available.

The *Tensorflow* module has a sub-module containing an API for the *Keras* library and provides a user-friendly interface to compose various deep learning models.

Sliding Window Method

In TS forecasting, future values have their basis on their past values. Because of this, the problem of forecasting is classified as a specific regression problem, an autoregression. In ML, regression is a type of supervised learning problem that refers to discrete outputs. In this context, supervised learning refers to adjusting a model by comparing the forecasted values with the real (observed) values. This adjustment is possible with the TS' past values and, as such, the data must be framed in that way.

In (Dietterich, 2002), a method to frame TSs as a supervised learning problem is presented. It is called the sliding window method (Figure 5) and it consists in transforming the data into several fixed-size combined input and output. Three variables are considered:

1. The stride s in which the windows are to be slid.
2. The number of timesteps of the inputs, i.e., the number of lagged values p in each window;
3. The number of timesteps of the outputs, i.e., the forecast horizons H in each window.

For the model's training phase, the width of the input $\boldsymbol{x}_t$ (Equation (1)) and the output are defined as,

$$\boldsymbol{y}_t = (x(j+1), \ldots, x(j+H)) \tag{1}$$

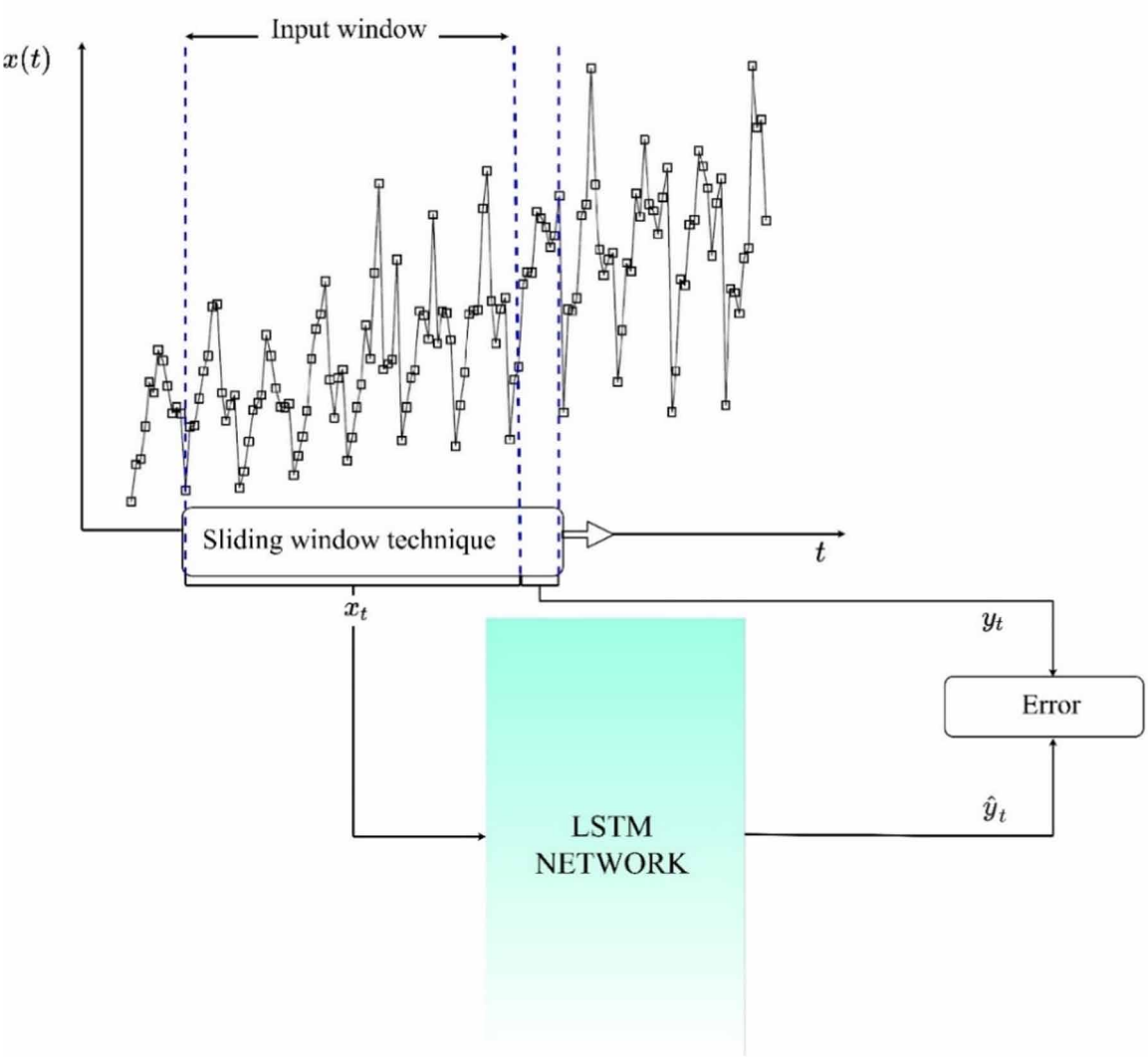
It is worth noting that H and p are sensitive parameters. While windows with few lagged values may prove to have insufficient information, large windows can increase the task's complexity and reduce the network's learning capabilities. In that respect, since no optimization was done, various p values are crudely picked for the half-hourly and the daily predictions. As for the stride s, a value of 1 is chosen, for all times, to maximize the number of samples that are provided to the model.

Model Configuration

For all instances, the networks are trained for 30 epochs and are composed by two layers with the following configuration:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The sliding window technique. Adapted from: (Paoli et al., 2010)



1. The input layer, which accepts windows of size p for every feature (in this case, only the variable of interest is used);
2. The LSTM layer has the hyperparameter *units* fixed at H. This parameter refers to the dimensionality of $\boldsymbol{h}_t$ that will be outputted. The other parameters use the default as defined by Keras[4] and no optimizations were done since the results are only of illustrative nature. As such, the activation functions (output and recurrent) are the same as in Figure 1.

Finally, for training purposes, the network uses MSE (Equation (17)) as the loss function and ADAM as ADAM as the gradient descent optimizer (Kingma & Ba, 2015).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Prophet

The model implemented in Prophet has two components: linear trend (Equation 12) and additive seasonality (Equation 13). Both terms of the model were adapted using the default parameters except the one related with the strength of seasonal fluctuations (seasonality_prior_scale=0.7) and the one that configures the trend's flexibility (changepoint_prior_scale). For the half-hourly TS, the latter was the only parameter changed (to 0.001), as its default value (0.05) is used in the daily example. The changepoint_prior_scale is the regularizer of the effects that changepoints cause in the TS. The remaining default parameters are the seasonality_mode (additive), the types weekly and daily for half-hourly TS seasonalities and the number of terms of the Fourier series (K=3 for both seasonalities).

Prophet needs to make a future data frame based upon parameters indicating the frequency and periods. Because of this, two TSs were required. For the daily TS, the forecasting horizon (the parameter period) is 28 days. In the other, the horizon is also 28 days, but it is also necessary to define the frequency of the predictions (freq='30min'). This information was added to the model after training the model.

RESULTS

This section presents the results of applying the two approaches to the case study presented above. Note that the main objective of this work is to explain the methods behind each approach and not to optimize the models to obtain the best possible results. Therefore, the results of each approach are not compared to each other and are only illustrative.

Daily Predictions

LSTM

For the daily predictions ($H = 28$), the LSTM model could not learn any seasonality component. It looks as if only the trend was captured and, because of that, unskillful forecasts are obtained. Values of p equal to 1, 4 and 7 were tested, and the results were very similar. Theoretically, using weekly windows (Figure 6) should have helped the network capture the weekly seasonality, but the results were equivalent or even a bit worse Table 1. A probable cause for this is the lack of available training samples as ANNs, such as the LSTM, are meant to be trained on large volumes of data.

Prophet

The forecast presented in Figure 7 shows that Prophet can capture weekly seasonality. The performance results are displayed in Table 2.

Prophet underestimated the number of clients who visited the store. The forecasts show a slightly decreasing trend possibly caused by number of clients visiting the store was above the usual during the first days in the training dataset. There are no changepoints and the weekly seasonality component is evinced in Figure 7.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Daily LSTM predictions for the last 28 days using the sliding window approach with $p = 7$

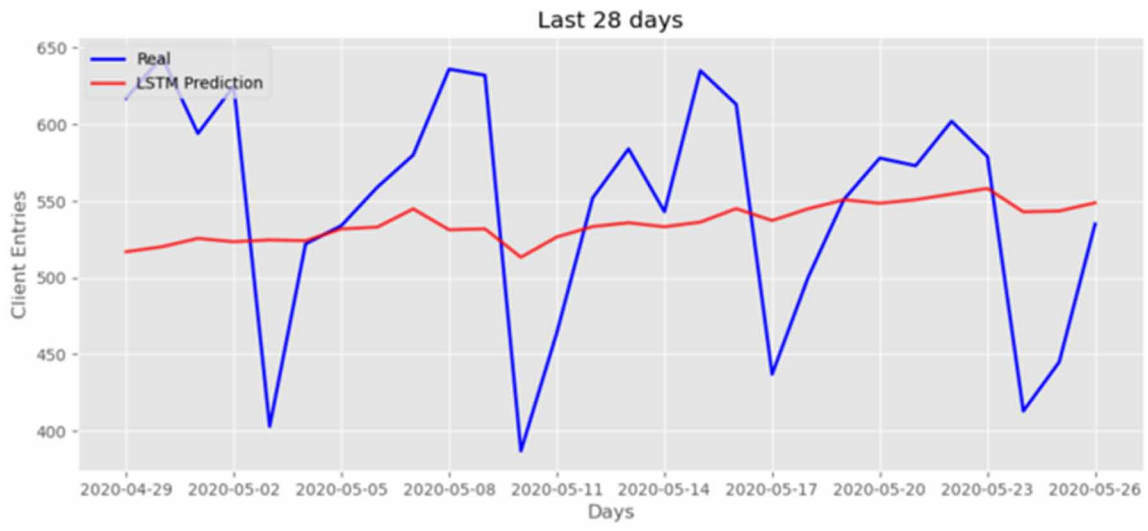
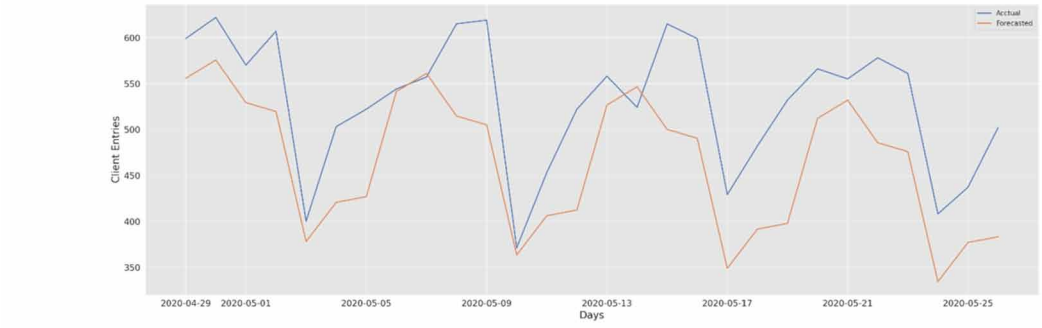


Table 1. Metrics for the daily LSTM prediction. Windows of one day offer best results

Daily	RMSE	MAE	MAPE(%)
$p = 1$ (a day)	**74.280**	**58.602**	**11.776**
$p = 4$ (four days)	75.503	60.510	11.848
$p = 7$ (a week)	75.410	61.795	11.883

Figure 7. Daily Prophet predictions and seasonality component of the last 28 days



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Metrics for the daily Prophet prediction

Prophet	RMSE	MAE	MAPE(%)
Daily	77.650	67.641	12.708

Half-Hourly Predictions

LSTM

In the half-hourly TS ($H = 700$), the objective of capturing weekly seasonality remains, yet, it is logical to consider daily seasonality as well. Multiple values of p up to 175 (weekly windows) were tested. The best results were obtained with $p = 50$ (Figure 8) as presented in Table 3. In contrast, Figure 9 shows the resulting predictions when the sliding window does not consider any seasonal period ($p = 10$).

Prophet

Prophet's predictions are displayed in Figure 10 and the performance results in Table 4.

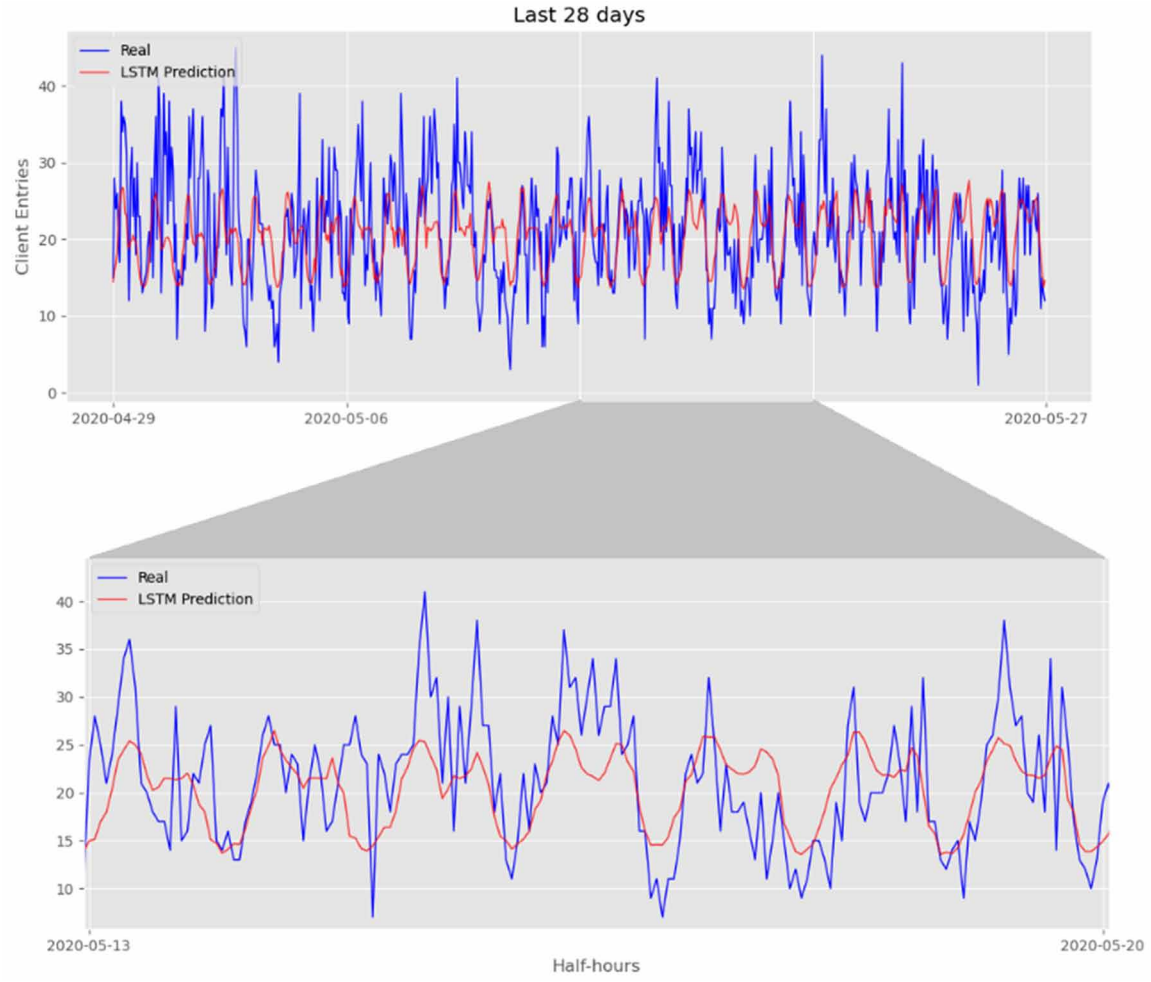
The gaps in the chart displaying the TS and the forecast correspond to the closing time. These gaps are not present in the charts in the previous subsection, because LSTM handles the data as a sequence of values regardless of the timestamps. The captured daily seasonality is represented in Figure 11, which indicates a peak in the number of clients between 11 and 14 hours. These results can be compared with Figure 3.

Table 3. Metrics for the half-hourly LSTM prediction. Windows of two days offer best results

Half-hourly	RMSE	MAE	MAPE(%)
$p = 10$ (five hours)	7.433	5.920	36.390
$p = 25$ (a day)	6.845	5.407	32.254
$p = 50$ (two days)	**6.450**	**5.097**	**30.101**
$p = 100$ (four days)	6.507	5.184	30.838
$p = 175$ (a week)	7.269	5.894	37.983

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Half-hourly LSTM predictions of the last 28 days with $\mathbf{p} = 50$ *(third week zoomed in)*



DISCUSSION

For the daily TS, it is evident that the sliding window configuration that offered more training samples to the LSTM model is the one that had better results. Nevertheless, results were not satisfactory as it is shown that the TS patterns were not captured. For small-size TSs such as this one, where DL approaches are unlikely to capture the seasonality, preprocessing methods to remove this component can improve performance.

For the half-hourly predictions, there are improvements in capturing seasonality. The lowest p values did not give the best performances but, surprisingly, neither a value of $p = 175$, which could have helped with the task of capturing the weekly seasonality.

With these results, it is possible to conclude that when modeling TS forecasting with LSTMs an equilibrium between p and the number of training samples should be achieved, especially for short TSs. Two additional concerns are valid when dealing with p:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Half-hourly LSTM predictions of the last 28 days with $p = 10$ *(third week zoomed in). As expected, five hourly windows are not suited for capturing daily seasonality*

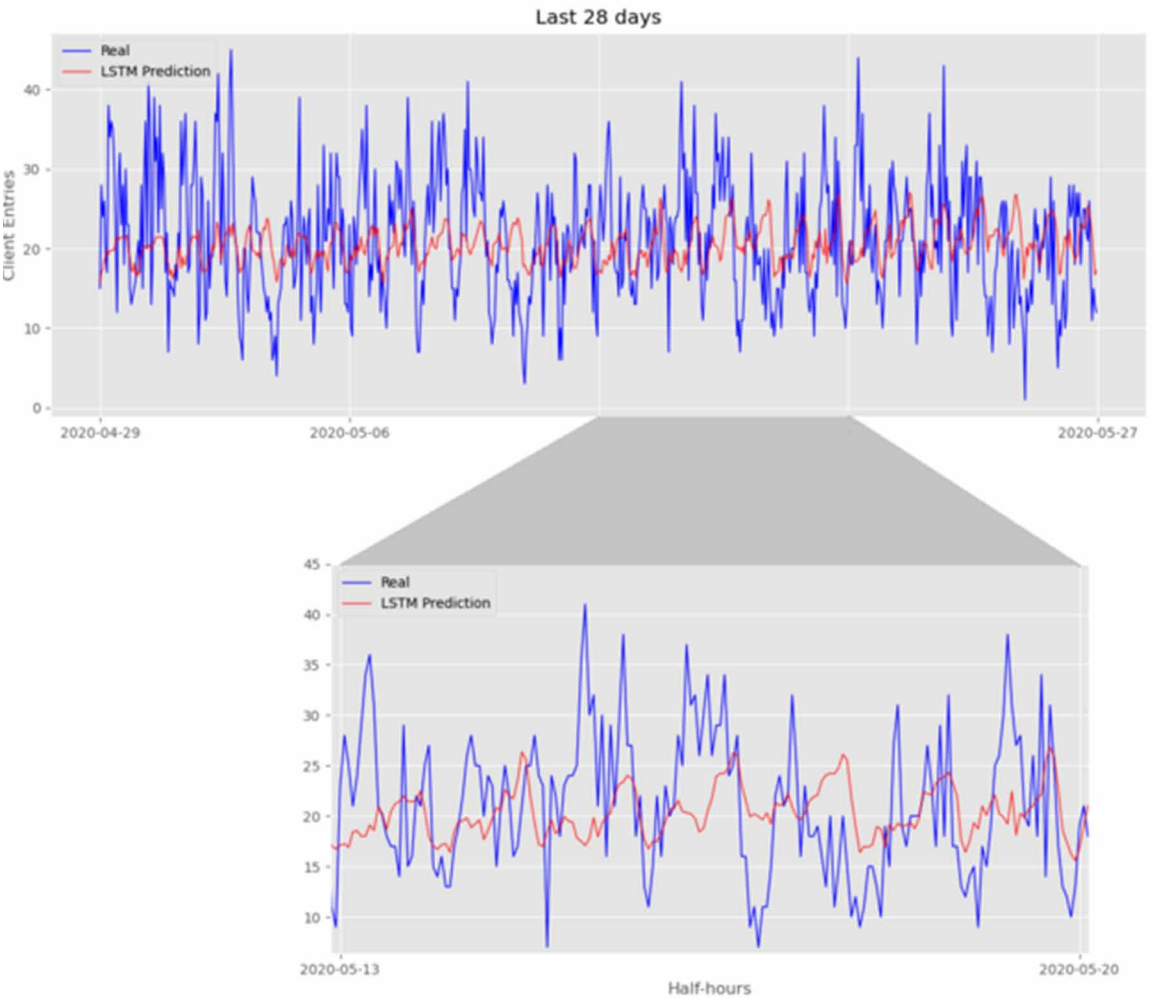
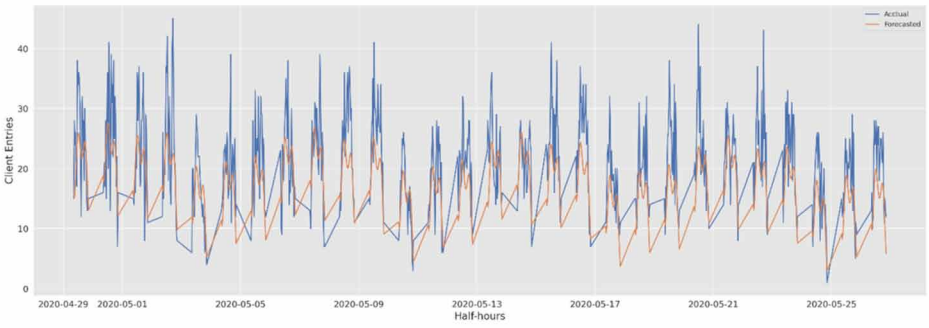


Figure 10. Prophet's half-hourly forecast against the real values



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

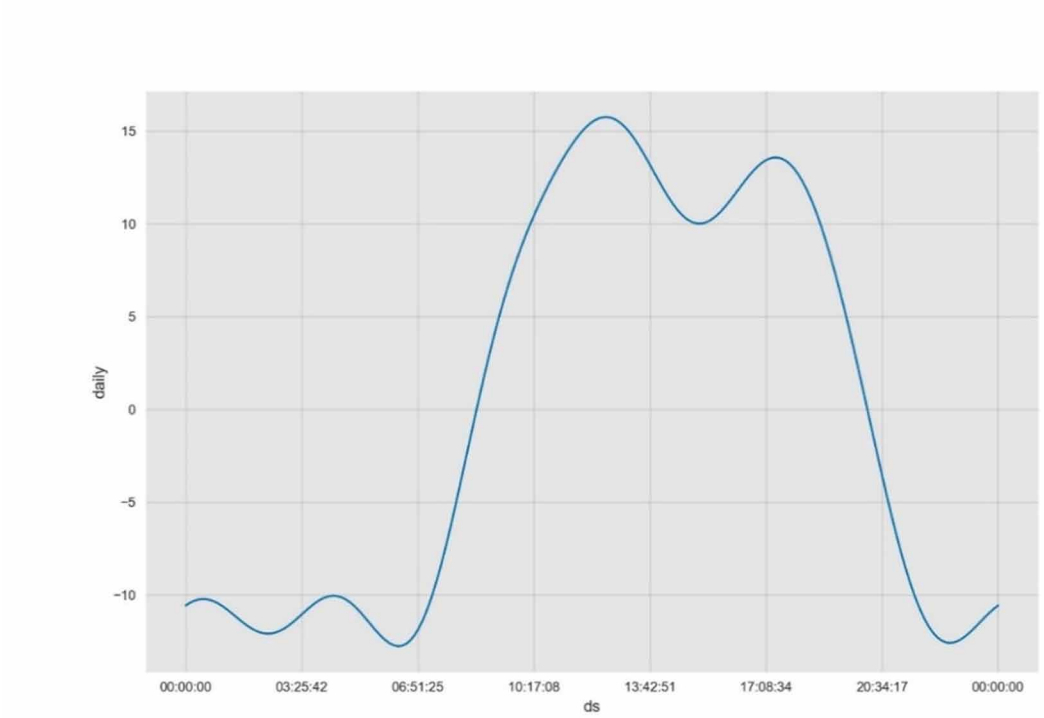
Table 4. Metrics for the half-hourly Prophet prediction

Prophet	RMSE	MAE	MAPE %
Half-hourly	6.512	5.157	24.687

1. The value should be small to lower the network's training complexity. Compared to $p = 50$, the model shows worse results when $p = 100$, meaning that larger windows do not necessarily mean better learning. This being related to the fact that, despite LSTMs having mitigated the problems of the traditional RNN, they still have limitations when maintaining lengthier long-term dependencies. For a shallow architecture as the one employed, limitations may be accentuated;
2. It should also consider the existing seasonality periods to allow the model to learn the seasonality components that may exist in the data (Figure 9).

Prophet seeks to automate the forecasting process as much as possible and, for that purpose, many of its parameters are already preconfigured. Prophet activates weekly seasonality when there are 2 weeks or more of historical data and the gap between dates in historical data is less than 7 days, and activates daily seasonality if there are 2 days or more in historical data and the gap between dates is less than 1 day. Most of the parameters used the default configuration, except the changepoint_prior_scale on the

Figure 11. Prophet's daily seasonality



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

half-hourly forecast, which allows less flexibility in the trend. Due to the higher temporal resolution and to improve the results, some changes in the parameters occurred on the half-hourly predictions.

Prophet proved to be an algorithm that fits these conditions and can be applied to TS along with several adjustments that may improve its accuracy. On the other hand, despite showing worse results, the LSTM model shows potential as the presented architecture is shallow. More layers can be stacked to the present model, enhancing its learning capability and possibly its predictive power. Another advantage of ANNs models is that various attributes can be included in the model without constraints. However, a disadvantage is that it requires a larger dataset to capture the TS characteristics (this is obvious in the daily case study). For the LSTM approach, it is worth noting that even though the TS does not show a significant trend, the seasonality component was not dealt separately. Deseasonalization methods are usually employed as a preprocessing task for DL models (Makridakis et al., 2018).

CONCLUSION

This work deals with time series forecasting in the retail business, using two recent approaches. The first uses LSTM, a general-purpose method based on recurrent neural networks, which can be applied to various problems where modeling temporal behavior is necessary. The second uses Prophet, an open-source library released by Facebook and specialized in business forecasting problems.

This chapter describes the methods used in each approach, explaining, at the same time, some of the main concepts and issues in TS forecasting. A simple case study is used to forecast the number of clients in retail stores per day and every 30 minutes, using real-world data. The concepts, problems and solutions are presented by example.

LSTM is a type of recurrent neuronal network that can learn and memorize early patterns in a data sequence. It was not designed for a specific use case, so it is more flexible and can be applied to a broader range of problems than Prophet, but it is also more challenging to use and the learning curve is longer. The application of LSTM to a specific use case requires in-depth knowledge of the method and on machine learning in general.

Prophet is a specialized library for business forecasting problems. The methods are already preconfigured and adapted to this application domain, and it is often possible to obtain acceptable results, even when using the default parameters. It is easier to use than applying LSTM models, but it is also less flexible because there are predefined choices by the library's developers. It also does not require in-depth knowledge of machine learning so that domain experts can use it more easily. Although, a basic understanding of programming in Python is assumed.

In future work, the detection, resolution and anticipation of outliers could also take part in improving forecasting performance. For these objectives, Prophet already has some means at hand. Another concept that can be relevant for retail companies is the incorporation of contextual data in forecasting problems, such as modeling the influence of weather conditions, promotional events, region-specific behaviors and many other variables.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

This work is partially funded by National Funds through the FCT (Foundation for Science and Technology) in the context of the projects UIDB/00127/2020 and POCI-01-0247-FEDER-039719. We also thank to SISQUAL (https://www.sisqualwfm.com/) for proposing the case study and for the valuable discussions on the subject.

REFERENCES

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), 1–41. doi:10.1016/j.heliyon.2018.e00938 PMID:30519653

Dietterich, T. G. (2002). Machine learning for sequential data: A review. *Lecture Notes in Computer Science, 2396*, 15–30.

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*(10), 2451–2471. doi:10.1162/089976600300015015 PMID:11032042

Hastie, T., & Tibshirani, R. (1987). Generalized Additive Models: Some Applications. *Journal of the American Statistical Association*, *82*(398), 371–386. doi:10.1080/01621459.1987.10478440

İşlek, İ., & Gunduz Oguducu, S. (2017). A Decision Support System for Demand Forecasting based on Classifier Ensemble. *Federated Conference on Computer Science and Information Systems (ACSIS)*, 13, 35–41. 10.15439/2017F224

Kingma, D. P., & Ba, J. (2015). Adam: {A} Method for Stochastic Optimization. In Y. Bengio & Y. LeCun (Eds.), *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015*, *Conference Track Proceedings*. https://arxiv.org/abs/1412.6980

Kolarik, T., & Rudorfer, G. (1994). Time series forecasting using neural networks. *Proceedings of the International Conference on APL : The Language and Its Applications, APL 1994*, 86–94. 10.1145/190271.190290

Maaß, D., Spruit, M., & de Waal, P. (2014). Improving short-term demand forecasting for short-lifecycle consumer products with data mining techniques. *Decision Analysis*, *1*(1), 4. doi:10.1186/2193-8636-1-4

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS One*, *13*(3), e0194889. Advance online publication. doi:10.1371/journal.pone.0194889 PMID:29584784

Paoli, C., Voyant, C., Muselli, M., & Nivet, M. L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, *84*(12), 2146–2160. doi:10.1016/j.solener.2010.08.011

Rosenblatt, F. (1957). The perceptron - A perceiving and recognizing automaton. *Issues (Chicago, Ill.)*, *Technical Report*, 85–460–1.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). A Comparison of ARIMA and LSTM in Forecasting Time Series. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 1394–1401. 10.1109/ICMLA.2018.00227

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 1–9. 10.1109/CVPR.2015.7298594

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, *72*(1), 37–45. doi: 10.1080/00031305.2017.1380080

Werbos, P. J. (1982). Applications of advances in nonlinear sensitivity analysis. In R. F. Drenick & F. Kozin (Eds.), *System Modeling and Optimization* (pp. 762–770). Springer Berlin Heidelberg. doi:10.1007/BFb0006203

KEY TERMS AND DEFINITIONS

Long Short-Term Memory (LSTM): A type of recurrent neural network that uses gating mechanisms to register long-term temporal dependencies in training data.

Machine Learning: A branch of artificial intelligence that studies computational algorithms that adapt their parameters accordingly to a given training dataset.

Neural Networks: Biological-inspired computational models that mimic the brain structure with simple processing units organized in layers that are highly connected.

Prophet: An open-source tool provided by Facebook for time series analysis and forecasting. It was designed to facilitate time series forecasting for business experts without deep knowledge in statistics or machine learning.

Recurrent Neural Networks: A class of neural networks that uses feedback loops to model temporal behavior in training data.

Retail: Process of selling services or goods using different channels to satisfy consumer needs on identified demands.

Time Series Forecasting: Deals with the estimation of future values using a model that was fit on observations collected over time (time series).

ENDNOTES

1 pandas.pydata.org
2 numpy.org
3 www.tensorflow.org
4 www.tensorflow.org/api_docs/python/tf/keras/layers/LSTM

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 12
General Model for Metrics Calculation and Behavior Prediction in the Manufacturing Industry:
An Automated Machine Learning Approach

Maria João Lopes
Bosch Termotecnologia SA, Portugal & University of Aveiro, Portugal

Eugénio M. Rocha
University of Aveiro, Portugal

Petia Georgieva Georgieva
https://orcid.org/0000-0002-6424-6590
University of Aveiro, Portugal

Nelson Ferreira
Bosch Termotecnologia SA, Portugal

ABSTRACT

In this chapter, a comprehensive description of a generic framework aimed at solving various predictive data-driven related use cases, occurring in the manufacturing industry, is provided. The framework is rooted in a general mathematical model so called queue directed graph (QDG). With the aid of QDGs and containerized microservices implementations, the typology of the system is analyzed, and real use cases are explained. The goal is for this framework to be able to be used with all use cases which fit in this typology. As an example, a data generation distribution model is proposed, the parameter stability and predictive robustness are studied, and automated machine learning approaches are discussed to predict the throughput time of products in a manufacturing production line just by knowing the processing time in their first stations.

DOI: 10.4018/978-1-7998-6985-6.ch012

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

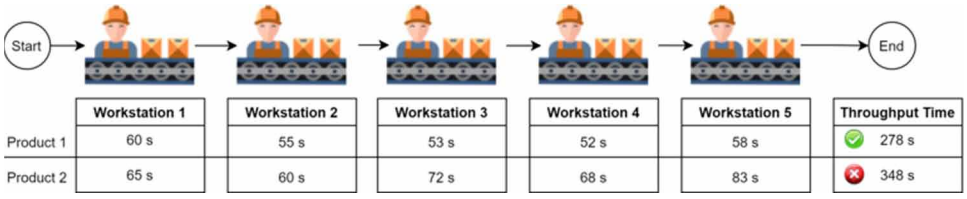
INTRODUCTION

A characteristic aim in the Manufacturing Industry is to optimize the operations in and around production lines (PL). For instance, simulation has been one of the most widely used techniques as shown in Russkikh (2020). Most of these sorts of problems, however, fit into a much broader category and with the right assumptions it is possible to create a model just generic enough to be able to solve a myriad of problems around this abstract entity, being it a PL, a value stream or a material delivery flow.

In order to understand the productive process, it is very common for organizations to resort to KPIs (Key Performance Indicators) and metrics. One of the most well-known is the *throughput time*. The throughput time describes the amount of time that it takes for a product to go through the production line. It can be calculated as the end processing minus the start processing time. Time based metrics are used extensively because they contain within other aspects of the process which would be hidden by other metrics. For example, if the throughput time exceeds some expected value, then it may reveal other underlying factors as the fact that not enough components had being fed to the PL, so the process will be starving at some point; the experience of the workers; a certain product family which is more complex to produce among others.

In Figure 1, a practical example of this calculation is featured. In this image, there is a schematical representation of a PL with a total of 5 workstations. The throughput time is calculated as the sum of all the *processing times* as duration of time the products spent in each workstation. Notice how there might be a trend already at the first 3 stations of the PL where Product 2 takes systematically more time than Product 1. This sort of trend is precisely what a prediction solution would like to catch and to see whether, from a small number of firsts stations of the PL, a significant trend in the throughput time can be inferred. With this information, one could classify the parts prematurely as efficient (such as Product 1) or non-efficient (as Product 2).

Figure 1. Schematic representation of the throughput time calculation in a production line



	Workstation 1	Workstation 2	Workstation 3	Workstation 4	Workstation 5	Throughput Time
Product 1	60 s	55 s	53 s	52 s	58 s	278 s
Product 2	65 s	60 s	72 s	68 s	83 s	348 s

The focus of this work is to construct a binary classifier where, given a certain effectiveness threshold plus the processing time at the first N stations, i.e., a small amount compared to the total size of the line, the solution is able to predict the outcome of the throughput time class.

In order to have an approach which would fit PLs in a broader sense, and even other types of systems with a similar topology, it is necessary to propose a generic model. The example that was seen above is actually an over-simplification of such manufacturing processes. For instance, it is considered that the throughput time is simply a sum over the different processing times in each workstation. However, there are actually waiting times in between the stations because they are not always available to start processing parts since they may be processing still the previous part. For this reason, the concept of queues will

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

also be introduced later on. Queues are entities where parts wait for their turn to be processed in a FIFO (First In, First Out) sequence.

Queues make the system more realistic yet there are other aspects to be taken into consideration such as reworks (repeating the process of a part more than once in the same station). In these scenarios, the part does not go back to the queue as this would be a waste of time for the process itself. However, this sort of changes to the usual topology must be taken into account.

A generic model allows one to calculate a myriad of different metrics which can occur at various levels. They can occur at the station level, for instance, the *processing time*, i.e., the time it take for the station to process one part. They can also occur at the transition level. One example is the *waiting time*, the time it takes for the part to end its processing in station N and to start its processing in station N+1. Finally, there are the so-called global metrics. The throughput time is an example of such a metric since it concerns the global entity, i.e., the PL. This classification is valid for metrics from the workstations point of view. There are also other types of metrics which can be calculated from the perspective of the part rather than that of the workstation. With this sort of generalization, this model would be valid not only for PLs, but for other systems which follow the same logic. These could be, for instance, trucks delivering goods or invoices flowing from supplier to customer.

The metrics described above can then either be presented through a dashboard or through an Early Warning System (EQS) which would notify the PL's Team Leaders of complications for a certain part. With this information, reactive measures could be taken immediately. The idea of doing the prediction for the throughput time metric arises precisely of the need of having enough time available to be able to take action in a meaningful timeframe. The main goal is also that the metrics can be reused and used for other higher-level applications, i.e., metrics calculation that are done on top of existing ones. This abstraction allows the calculation of other metrics in a much easier way since the baseline already contains processed information one can use.

In the next section, some use cases found in the literature will be reviewed, along with the Machine Learning (ML) techniques that can be leveraged to solve such issues. Afterwards, the focus will shift towards presenting a mathematical data model and its characteristics. Defining a precise model is quite relevant to ensure that all assumptions are fulfilled by the real data collected and that the model accounts for particular phenomena of the physical process, that in several models are simply not taken into account. Once this theoretical construct has been properly introduced, the robustness of the method will be tested through a series of experiments using simulated data. With these results as basis, the method is applied to a real PL. Subsequently, the authors dwell on the challenges of implementing such use cases in an organization and propose some suggestions to tackle them. The chapter ends with the conclusion and future research directions to be followed.

BACKGROUND

The usage of mathematical methods and Artificial Intelligence (AI) is becoming an increasingly studied subject in the academia in recent years. However, it can be quite challenging for executive and upper-level management to pinpoint actual usages which can translate this premise into actual value for their organizations. A more detailed look into the challenges that these entities are facing can be found near the end of the chapter.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Columbus (2019) presents different examples of the utilization of ML techniques to the added benefit of manufacturing corporations. These range from supply chain, to process or product design optimization. Even more qualitative areas such as Human Resources can benefit tremendously from these applications.

However, the real adoption rate is still reduced. According to Lee (2020), there are few articles that focus on direct applications on ML techniques at the enterprise level. The ones available focus mainly in the banking and finance sector. These are the sectors that are more heavily invested in these technologies. In industrial settings, the most widely known applications are generally reserved to the tuning of process parameters, as discussed in Chen (2017) and Weichert (2019). These can be applied in a wide range of processes such as Welding or Plastic Injection Molding. Still, the overall scenario in the entrepreneurial realm is that there are few actual deployments and most of them are still in a pilot phase.

Another important aspect mentioned in the literature is the process of algorithm selection to solve a concrete problem. This is an aspect the authors are deeply interested in tackling and one reason why the development of the framework at hand is being done using an agile approach. This way one can clearly define the characteristics that a problem must have in order to fit in a certain category. This is a widely known issue, that there is not one single algorithm that proves to be a good fit for all problems. This limitation is explored extensively in Pedro Domingos' book, The Master Algorithm (Domingos, 2017).

With this context in mind, one can still find some examples in the recent literature which address some areas which will be in focus during this chapter. One of them are Random Forests classifiers which are explored in Wu (2017). In this article, the authors compare the performance of Artificial Neural Networks, Support Vector Regression and Random Forests since in previous literature, there was a good performance of decision trees which are closely related method to Random Forests. Using Random Forests proved a reliable solution according to the author's conclusion. The method was able to predict accurately the wear of the tool in milling processes. Another completely different application of this algorithm will be explored in this work.

In more specific terms, this chapter covers the usage of a graph based, generic approach for solving a number of different problems. Here, the approach will be focused in the prediction of the throughput time. In Backus (2006), the focus was on predicting the cycle time of intermediate cycle times in the semi-conductor industry. One of the issues the researchers ran into was the influence of outliers in the predictive model. For this reason, the batches which exhibited a deviating behavior were excluded from the scope. The authors concluded that a real time application would be a desirable next step. These two questions, how to treat the outliers and how to use the model in real time are important focus points in this work.

This study will also cover the usage of Automated ML (AutoML) as a way of achieving high performance results. In Tušar (2017), overfitting is referred and conclusions state that it can be reduced by using tree ensembles, giving as an example the Random Forests model. However, AutoML is known for being intrinsically an overfitting strategy, since a heavy number of combinations and optimizations are ran in order to achieve the best possible results. It will be assessed whether this characteristic can be used to one's advantage later.

DATA MODELLING

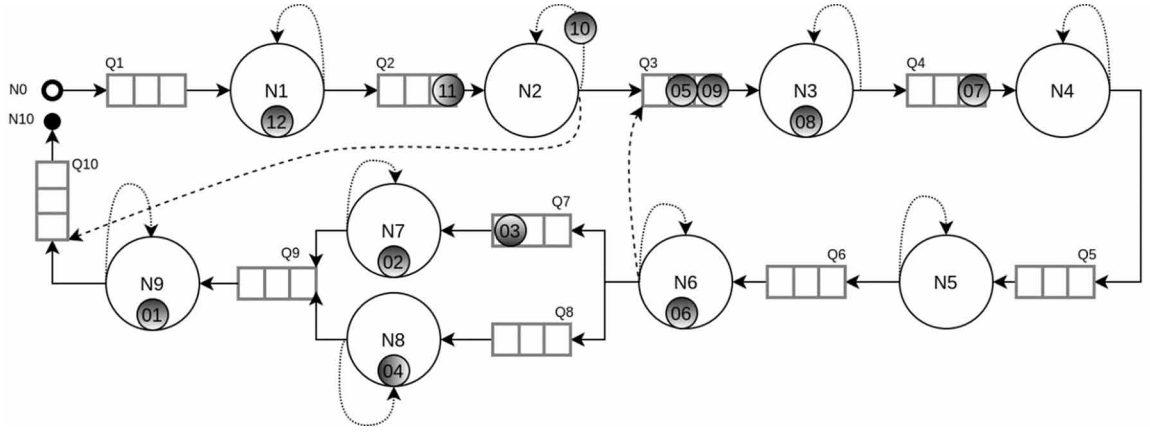
A production line (PL) may be modeled in quite different ways, depending on the goal of the study. Common ways to formally describe workflow processes include Petri net variants or activity diagram

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

variants. Even though they show similarities, fundamental differences turn their applications into PLs too complex or lacking relevant features, see Eshuis and Wieringa (2003). Several of these models assume that parts are produced or consumed by actors and do not incorporate (explicitly) the notion of layers of queues and their associated nodes, see Buck and Lee (1993).

For a better understand of the problem, a mathematical structure for modelling a graph with queues is introduced, allowing to incorporate the assumptions and metadata that is common in a PL. At its core, a graph is merely a set of nodes and arcs, which can easily be associated with physical entities (workstations, people or servers). The nodes are connected between one another by arcs, describing paths of product transitions. Nodes and arcs are the infrastructure but truly important are the products, the data that they have associated, and data operations that can be applied in each component of the graph.

Figure 2. Schematic representation of a production line, its stations and products flowing in it



In Figure 2, nodes (identified from N_1 to N_9) represent transformation stations in the PL, tokens (small dark circles with numbers from 01 to 12) represent parts or materials moving in the PL, queues (identified by Q_1 to Q_{10}) are intermediate locations where tokens wait for the respective station to process them, N_0 is the starting location of tokens, N_{10} is the final location of tokens, and direct arrows represent transitions and component interconnections. Dotted arcs represent the particular situation where a token is re-worked in the same node without any other token being worked in the station. Dashed arcs represent the situation where tokens do not follow the standard layout of the PL, jumping to a different station or being removed from the production area (e.g., anomaly detection and removal).

General Model for a Manufacturing Production Line

To have a precise definition for the problem, with similar representation as Figure 2, the mathematical notion of a Queue Directed Graph (QDG) is introduced as an ordered tuple $(N, Q, A, B, C, K, W, G, dt_B, dt_E)$ where:

1. N is a list whose elements are called nodes with $s_N + 1$ elements and indices in the set $i \in \{0, \ldots, s_N\}$;

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Q is a list of s_Q first-in/first-out queues (FIFO queues) with no capacity limit and indices in the set $j \in \{1,\ldots,s_Q\}, s_Q = s_N$;
3. A is a set with a minimum of s_N directed arcs as pairs of node-queue of the form (N_i, Q_j);
4. B is a set of s_N directed arcs as pairs of queue-node of the form (Q_j, N_i), which associates one queue to its corresponding node for all nodes in the graph except for N_0;
5. K is a list of $s_K + 1$ tokens with indices in the set $k \in \{0,\ldots,s_K\}$;
6. C is a matrix of integers where $c_{i,j}$ accounts for the number of tokens that have transverse the arc (N_i, Q_j);
7. W is a list of tuples with the form $W_k = (din_w, dout_w, N_w, K_w, P_w, G_w)$ with indices in the set $w \in \{0,\ldots,s_W\}$, where din_w is the timestamp when the token K_w enters the node N_w, $dout_w$ is the timestamp when token K_w leaves N_w, $din_{w+1} \geq din_w$, P_w is the duration of time that the token K_w was in N_w before leaving, and G_w is a set of words used for Root Cause Analysis (which may be empty) associated with the processing time;
8. G is a set of s_G tags (words) that are used to classify 'situations' (this set can be empty);
9. The operations on the QDG are performed in the interval $[dt_B, dt_E]$;
10. Tokens do not spend time moving on arcs (i.e. they move instantaneously);
11. The node N_0 is the starting node and node N_{s_N} is the ending node;
12. Tokens can only enter through N_0, can leave from N_{s_N}, and cannot enter the graph more than once.

Note that, because of assumptions 3 and 4, there are no arcs from nodes to nodes or queues to queues. Moreover, because in assumption 3 the cardinality of A is not fixed, the graph transition may not be deterministic (i.e. tokens may follow different paths on exiting a node). Each token defines a path in the QDG, which indirectly is coded in the matrix C, from where one may associate to each arc a transition probability. This opens the possibility to apply Stochastic and Probabilistic Methods to QDG (e.g. Markov Chains or Bayesian Networks). Furthermore, the graph in Figure 2 is not a QDG, since it has node loops (i.e. fails assumption 12), but may be solved through operation E_1.

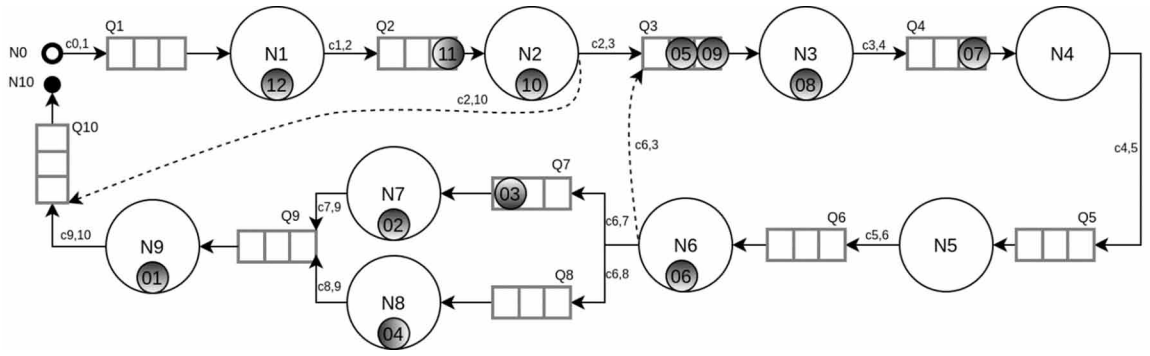
(E_1) As stated above, in practical situations a token (i.e. component/material/product) can be reworked at the same node (i.e. station in the PL) without any other token entering the node. This situation is not directly compatible with a QDG, but can be treated by supposing that two consecutive entries in W as $(din_w, dout_w, N_w, K_w, P_w, G_w)$ and $(din_{w+1}, dout_{w+1}, N_w, K_w, P_{w+1}, G_{w+1})$ are replaced by the entry $(din_w, dout_{w+1}, N_w, K_w, P_{new}, G_{new})$, where $P_{new} = dout_{w+1} - din_w + P_w$ and $G_{new} = G_w \cup G_{w+1}$.

(E_2) Human operations may break predefined rules as the FIFO order of a queue. In such case and to comply to a QDG, the situation is reported and the involved tokens are completely removed from the data set for this study.

From now on, the authors assume that the above situations are treated by the operations $E_1 - E_2$, so one can have a well-defined QDG as in Figure 3.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Schematic representation of a Queue Directed Graph (QDG) with the associated metrics calculations



This mathematical formulation is quite general and can be applied to many real scenarios, which will be explored in future publications.

The QDG structure, along with an integration into a technology infrastructure, allows to define metric levels. Metric levels are increasing self-contained layers where metrics (e.g. KPIs) are defined from metrics of previous levels. In the metric layer 0 sits a QDG. The authors now provide an example of two metric layers.

Metrics level 1: Considering the existence of a well-defined $QDG = (N, Q, A, B, C, K, W, dt_B, dt_E)$, one needs to define metrics related with:

- Tokens last processing locations: A list LP with the size s_K that in the position k (associated with the token of index k) holds the pair $(dout_k, N_k)$ where $dout_k$ is the last timestamp of departure of the token from the node N_k. The list starts empty;
- Node idle times: A list IT with the size s_W that holds the idle time duration (a float) for each entry W. The list starts empty.
- Node accumulated processing: A counter with a value smaller than s_K which enables to keep track of how many tokens have been processed for the desired timeframe e.g. the current shift. The counter starts as zero and may be reset according to the desired metric, e.g. daily or at shift start.
- Tokens throughput time: An attribute for each token k which measures the time elapsed in $[dt_B, dt_E]$ spend to go from the start node to the end node.

Then, for each $w \in \{0, \ldots, s_W\}$, so having $W_k = (din_w, dout_w, N_w, K_w, P_w, G_w)$ and $k = index(K_w)$, the above metrics can be calculated as: (a) $IT_w = LP_{k,0} - din_w$; (b) $LP_w = (dout_w, N_w)$.

Metrics level 2: From the metrics obtained in the previous level, one may calculate for each element (i.e. nodes, queues, tokens) but also globally for the PL, the following metrics: (a) node idle time (from IT_w); (b) node processing time (from P_w); (c) node locking time (as the sum of idle time plus processing time); (d) node starving time (as the difference between the total study duration and the total locking time); (e) node throughput (as the number of token leaving a node per time unit); (f) queue permanence (as the number of token in the queue per time unit), etc..

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The authors would like to remind the reader to the fact that, although this mathematical approach is quite relevant for the purpose of this study, it is still general enough to accommodate many other problems, e.g. bottleneck detection and prediction or predictive quality control. This gains particular relevance when one considers the usefulness of level 2 metrics and higher.

Workstation Processing Times Data Generator

Although the proposed approach will be tested against data from a real manufacturing PL, to be able to proceed with several experiments regarding parameter stability and prediction robustness, fundamental to benchmark the chosen problem and also the machine learning approach, it is required to have a mechanism to generate mathematical valid data from the real data. In the literature, the authors were able to find works regarding simulation and data generation of data for a PL as Centomo (2018). However, the mentioned approaches made too many simplifications or are not easily adapted to the problem under study. Additionally, the purpose here is to find a mathematical model as accurate as possible to this situation.

In this work, in order to generate data under valid assumptions, a QDG model is used where artificial data is obtained by simulation by using two generators to construct the values of W_k (see assumption 7 on the definition of a QDG): (a) a node-to-node path generator for each token; and (b) a processing times generator for each node. The path generator is simply obtained from real data by constructing a count transition matrix node-to-node, from which we obtain the probability of a token leaving a node *i* and moving into a node *j*. The processing times generator is not so straight forward and needs to be adapted to the real data, as much as possible.

A common approximation for the processing times (PT) distribution is to use a normal distribution. However, such is quite imprecise, since processing time histograms of real data show that the data is not symmetric, does not have negative values or even values below some positive threshold, and outliers tend to follow a different distribution. Such implies that a more realistic distribution function for the generator will be a composite probability distribution, see Oosterbaan (2019).

As expected, measures between probability distributions play a significant role in the quest to find a good generator. There are many variants such as the Hellinger coefficient, Jeffreys distance, Chernoff coefficient, directed divergence, and Kullback–Leibler divergence (relative entropy). Using the relative entropy, the probability density function (p.d.f.) of processing times of real data and the p.d.f. associated to a processing times generator may be compared to measure how near they are to each other. Then, by experts' intuition and tentative and error, several processing times generators may be created and tested to find the best candidate. The generator will depend on parameters that take into account the variability and specificity of each workstation, which can be found by traditional grid fitting methods on real data. With the above approach, the generator proposed in this work is defined as follows:

Let $U(a,b)$ and $LN(\mu,\sigma)$ denote (respectively) a uniform distribution defined in $[a,b] \subset \mathbb{R}$ and a log-normal distribution with location parameter μ and shape parameter σ. Define the constants $0 < a_1 < b_1, a_2 < b_2, 0 < a_3 < b_3, 0 \le a_4 < b_4 < 1$, the maximum percentage of outliers $r \in [0,1[$, and for each processing location $i \in \mathbb{N}$:

- $w_i \in U(a_1, b_1)$ as the lower production barrier;
- $¶_i \in U(w_i + a_1, w_i + b_1)$ as the upper outlier barrier;

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- $¼_i \in U\left(w_i + a_2, w_i + b_2\right)$ as the location parameter (e.g. mean);
- $\tilde{A}_i \in U\left(a_3, b_3\right)$ as the shape parameter (e.g. standard deviation);
- $^3{}_i \in U\left(a_4, b_4\right)$ as the outlier selection parameter;
- the generator function for the processing times as

In this definition, P_i behaves as a random variable with bounded support that generates the processing times at the node *i*. The first branch generates abnormal processing times and the second generates the normal ones. The distribution has support above w_i because it is not realistic to assume that a machine or an operator can shrink the processing time to zero, without a lower limit, just by adding more effort without scenario changes (e.g. increasing hardware or human resources). As expected, the above parameters can be precisely found for a given data set but also allow to model different data behaviors on each processing location.

As is predicted by the famous Central Limit theorem, the distribution of the sum of (a big enough) number of processing times at different locations will approach a Normal distribution, which is reinforced by the fact that some products may leave the PL earlier (e.g. defect detection at a station), thus contributing to the symmetry of the distribution of the throughput time as a sum of the distributions of the processing times at each workstation.

THROUGHPUT PREDICTION ON SIMULATED DATA

The main goal of this section is to describe the various tests that were conducted in order to find out whether it would be reasonable for the proposed framework to predict in real time the efficiency of a given token. In practice, this means that the authors have started with a feature of the graph such as the throughput time (e.g. the time it takes for a token to navigate the graph from start to end). Based on a given percentile one calculates a label based on it: 0 (not efficient) or 1 (efficient). What will be assessed is whether one can accurately predict this label by using incomplete information about the feature at hand. For this use case, the first nodes and transitions durations will be used in order to predict the label's outcome. If the authors succeed, then one can predict a token's performance using just the data from the first nodes. Since this is a graph, its length is completely irrelevant since this applies to all use cases. What might prove meaningful is the ratio between the nodes used for prediction and the number of nodes in the graph.

Parameter Stability and Predictive Robustness

To have a valid industrial solution, it is of major importance to measure its stability to a change in the conditions. This will have a consequent impact in the robustness of the algorithmic solutions used. The following experiments have been conducted using simulated data as described in the *Workstation Processing Times Data Generator* subsection, in order to have a data set with the intended characteristics. Note that the data sets generated with the distribution proposed are similar (in a precise mathematical sense) to a data set with data of a real production line. A graph with a typology of 9 nodes has been used.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Results for experiment B, where the accuracy of the pipeline was tested with different levels of information and against different percentiles

	Percentile = 60%		Percentile = 80%	
	Single Features	**Single + Enriched Features**	**Single Features**	**Single + Enriched Features**
2 out of 9 nodes	58.18%	65.45%	79.39%	78.18%
3 out of 9 nodes	66.66%	70.30%	78.78%	81.21%

Experiment A – Class of machine learning algorithms: The first step to have a controlled study about parameter stability and predictive robustness, allowing the use of hyperparameter and pipeline optimization, is to define a class of techniques to optimize and test. As described in the introduction, the Machine Learning (ML) problem to be solved may be reduced to a binary classification.

There are many techniques in the literature that may be applied to binary classification with proved success such as: (a) Naïve Bayes, (b) Logistic regression and variants; (c) k-Nearest neighbor with Euclidean or Mahalanobis distances; (c) Support vector machines and variants with kernel functions; (d) Variants of decision trees as bagging decision trees, boosted decision trees, or random forests; (e) Gaussian mix models; and (f) Voting classification.

After testing each of the above techniques in the data sets of this work with sparse grid hyperparameter optimization (whose descriptions are omitted), the conclusion is that all techniques perform relatively well. However, the variants of decision trees are the best ones when compared to the standard metrics obtained from the confusion matrix. For such reasons, from now on, the experiments use techniques such as these and were implemented in standard Python libraries such as Scikit-learn and TPOT, for automated machine learning (autoML).

Although the percentile threshold predefined will characterize if the classes are balanced or not (i.e. being near 75%, mean that it exists 75% of zeros and 25% of ones), the number of false positives and false negatives obtained were quite low, so the focus will not be on F1 but rather on the accuracy metric.

Experiment B – Number of nodes and percentile influence: This experiment intends to understand if there is a dependence between the accuracy metric results obtained by a Random Forest Classifier (RFC) and the changing number of features (i.e., number of nodes used for prediction). The label is obtained from the percentile of the total time that tokens are in the graph (i.e., going to all the nodes and queues of their complete paths).

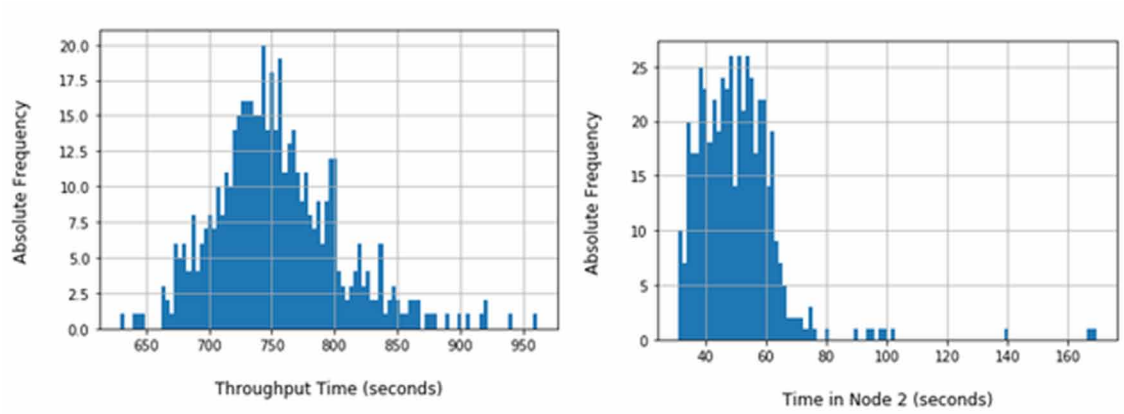
Using the standard Scaler in conjunction with RFC with 100 estimators, and without any further fine tuning to the pipeline, the accuracy results described in Table 1 were achieved. Two thirds of the data were used for training whilst the remaining were used for testing.

Here, the training algorithm was run for two different percentiles to assess the impact of the percentile in the overall outcome. Additionally, the authors have also ran a single data set and a data set with enriched features. Single features are the durations of nodes and transitions (time spent in the queue), while enriched features include some of these durations summed together. Since these aid the algorithm to find correlations, one can see a significant increase in the accuracy when the enrichment is added. The only exception seems to have occurred for the 2 out of 9 nodes experiment with percentile 80%. This deviation may be due to inherent characteristics of the data set.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

One can here observe already very satisfying levels of prediction which indicate that the data set is quite stable. This can be confirmed by analyzing the histogram of the throughput time from which the percentile is calculated, see Figure 4 (left). One verifies that the percentile introduces some perturbation in the accuracy of the classifier thus making it perform better for higher percentiles.

Figure 4 (left) and Figure 4 (right). On the left, histogram from the metric from which the percentile is calculated. On the right, histogram of the processing time in the second node of the graph. The number of bins was adjusted to provide enough granularity.



Experiment C – Outlier influence: In this experiment, the outlier variation was set to an uniform distribution with parameters 0 and 0.5 i.e. there was a significant reduction of the outlier variation in comparison to Experiment B.

Table 2. Results for Experiment C where the accuracy of the pipeline is tested against a more reduced level of outliers

	Percentile = 80%	
	Single Features	**Single + Enriched Features**
2 out of 9 nodes	78.79% (-0.6%)	79.39% (+1.21%)
3 out of 9 nodes	81.81% (+3.03%)	83.03% (+1.82%)

The 60% percentile was dropped as this is not a sufficiently realistic percentile level for real applications.

Compared with Experiment B (see Table 1), one can see that in the first row of Table 2, there is not much improvement of the classifier accuracy. Most probably, this is due to very little data being used, which translates into higher deviations.

In the second row of Table 2, there was a slight increase in accuracy for both single and enriched features compared to Experiment B. The underlying reason is that the enriched features tend to soften the variability that is introduced by the outliers.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Results for experiment D where the accuracy of the pipeline is compared among different data sets

		Percentile = 80%	
		Single Features	**Single + Enriched Features**
Data set 0	2 out of 9 nodes	73.94%	73.94%
	3 out of 9 nodes	76.36%	79.40%
Data set 1	2 out of 9 nodes	76.36%	78.79%
	3 out of 9 nodes	84.24%	86.67%
Data set 2	2 out of 9 nodes	78.18%	78.79%
	3 out of 9 nodes	81.82%	81.82%
Data set 3	2 out of 9 nodes	78.18%	79.39%
	3 out of 9 nodes	80.61%	81.21%
Data set 4	2 out of 9 nodes	77.58%	79.40%
	3 out of 9 nodes	82.42%	83.03%

Experiment D – Data set variability: In the previous experiments the ML pipeline was tested in one data set only. In this experiment, 5 different simulated data sets are used, in order to study the robustness of the classifier. Outlier variability is the same as in the previous experiment.

Overall, the classifier robustness is within reasonable limits of variability. The interval between maximum and minimum accuracy increases as more nodes are used (see Table 4). This makes sense since the number of features is also being increased and thus variability increases.

Again, the smoothing effect provided by the enrichment is an advantage for classifier's performance as one can assess by comparing single versus singles plus enriched features. This holds even for this larger sample as expected.

Experiment E – Local and global perturbations (location parameter): Up to now, the simulator was using a range of parameters for the log normal distribution for the duration in each transition and node. In this experiment a deviation will be introduced in Node 2. The goal is to test how this deviation affects the overall classifier performance, which is tied to the end metric, the Throughput Time. Here, the parameter μ of the log normal (i.e. the location parameter) is shifted by an additional 2 units. These units are added to a parameter which is generated from an uniform distribution.

In overall terms, the classifier's performance is equivalent to that of Experiment D. Regarding classifier stability, there is a trend for an increased amplitude in the 2 out of 9 nodes scenario. Yet on the 3 out of 9 nodes, the amplitude shows a decreasing trend which translates in better stability even when one tends to the fact that a significant deviation was introduced.

Table 4. Statistics for the results in Table 3

		Minimum	**Mean**	**Maximum**	**Δ (Max-Min)**
2 out of 9 nodes	Single Features	73.94%	76.84%	78.18%	4.24%
	Single + Enriched Features	73.94%	78.06%	79.40%	5.46%
3 out of 9 nodes	Single Features	76.36%	81.09%	84.24%	7.88%
	Single + Enriched Features	79.40%	82.42%	86.67%	7.27%

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5 (left) and Figure 5 (right). On the left, the histogram from the metric from which the percentile is calculated. On the right, histogram of the duration in Node 2 where the introduced deviation is visible.

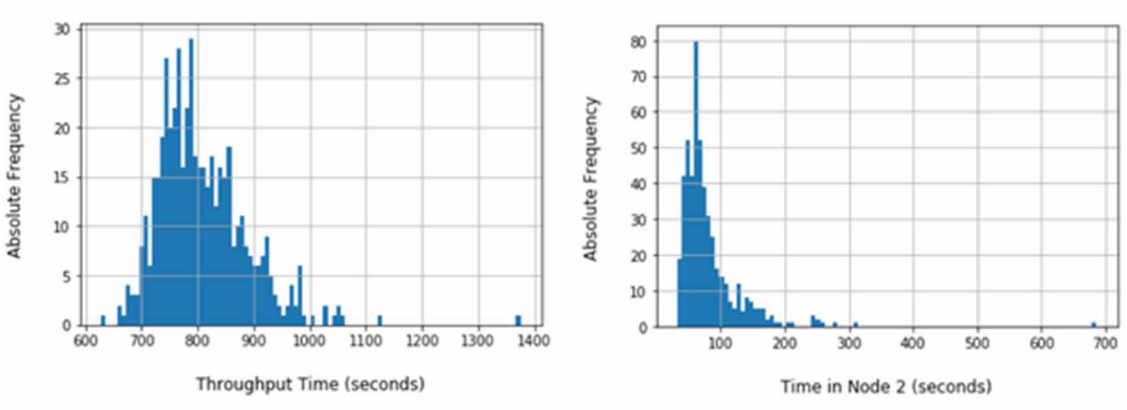


Table 5. Results for experiment E where the accuracy of the pipeline is compared among different data sets with the introduced perturbation

		Percentile = 80%	
		Single Features	**Single + Enriched Features**
Data set 0	2 out of 9 nodes	79.40%	81.82%
	3 out of 9 nodes	83.03%	82.42%
Data set 1	2 out of 9 nodes	80.61%	81.82%
	3 out of 9 nodes	81.21%	81.21%
Data set 2	2 out of 9 nodes	81.82%	78.18%
	3 out of 9 nodes	86.06%	83.64%
Data set 3	2 out of 9 nodes	80.61%	81.21%
	3 out of 9 nodes	82.42%	81.21%
Data set 4	2 out of 9 nodes	76.97%	78.79%
	3 out of 9 nodes	80.00%	80.00%

Table 6. Statistics for the results in Table 5

		Minimum	**Mean**	**Maximum**	**Δ (Max-Min)**
2 out of 9 nodes	Single Features	76.97%	79.88%	81.82%	4.85% (+0.61%)
	Single + Enriched Features	78.19%	80.36%	81.82%	8.63% (+3.17%)
3 out of 9 nodes	Single Features	80.00%	82.55%	86.06%	6.06% (-1.82%)
	Single + Enriched Features	80.00%	81.70%	83.64%	3.64% (-3.63%)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6 (left) and Figure 6 (right). On the left, histogram from the metric from which the percentile is calculated. The effect of the deviation is quite drastic. On the right, the histogram of the duration in Node 2 where the introduced deviation is also clearly visible.

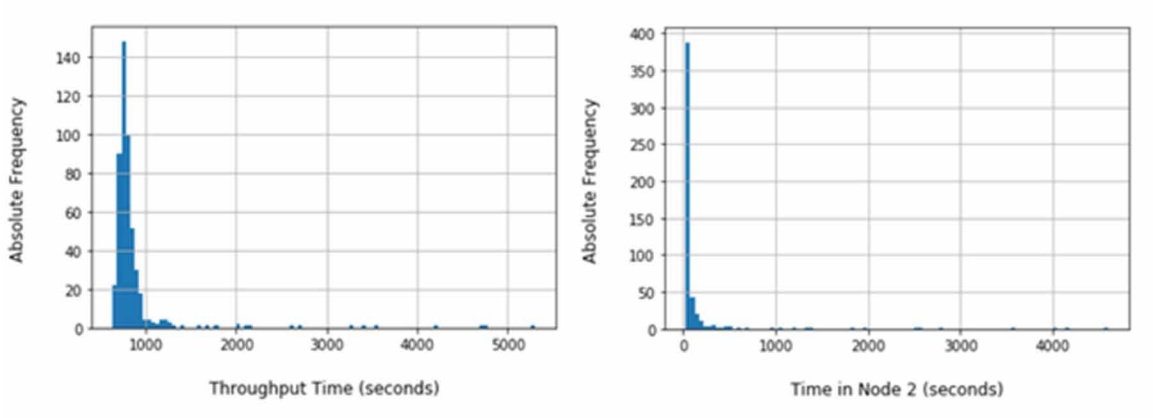


Table 7. Results for experiment F, where accuracy of the pipeline between different data sets was compared

		Percentile = 80%	
		Single Features	**Single + Enriched Features**
Data set 0	2 out of 9 nodes	86.06%	84.24%
	3 out of 9 nodes	86.06%	85.46%
Data set 1	2 out of 9 nodes	89.09%	90.30%
	3 out of 9 nodes	88.49%	89.09%
Data set 2	2 out of 9 nodes	85.46%	84.24%
	3 out of 9 nodes	86.06%	87.88%
Data set 3	2 out of 9 nodes	87.88%	88.48%
	3 out of 9 nodes	90.30%	90.30%
Data set 4	2 out of 9 nodes	85.45%	87.27%
	3 out of 9 nodes	89.09%	88.49%

Table 8. Statistics for the results in Table 7

		Minimum	**Mean**	**Maximum**	**Δ (Max-Min)**
2 out of 9 nodes	Single Features	85.45%	86.79%	89.09%	3.64% (+0.6%)
	Single + Enriched Features	84.24%	86.91%	90.30%	6.06% (-0.6%)
3 out of 9 nodes	Single Features	86.06%	88.00%	90.30%	4.24% (+3.64%)
	Single + Enriched Features	85.46%	88.24%	90.30%	4.84% (+2.43%)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9. Results obtained using TPOT for 5 different data sets with percentile set at 75%

	Accuracy Test Set	Accuracy Training Set	Pipeline
Data Set 0	100%	99.87%	GradientBoostingClassifier(learning_rate=0.5, max_depth=8, max_features=0.25, min_samples_leaf=19, min_samples_split=19, n_estimators=100, subsample=0.85)
Data Set 1	100%	99.93%	XGBClassifier(learning_rate=0.001, max_depth=8, min_child_weight=6, n_estimators=100, n_jobs=1, subsample=0.55, verbosity=0)
Data Set 2	100%	99.80%	XGBClassifier(learning_rate=0.01, max_depth=2, min_child_weight=2, n_estimators=100, n_jobs=1, subsample=0.3, verbosity=0)
Data Set 3	99.8%	100%	DecisionTreeClassifier(criterion="entropy", max_depth=4, min_samples_leaf=6, min_samples_split=13)
Data Set 4	100%	99.73%	StackingEstimator(estimator=GradientBoostingClassifier(learning_rate=0.01, max_depth=7, max_features=0.2, min_samples_leaf=8, min_samples_split=5, n_estimators=100, subsample=0.8)), GaussianNB())

Experiment F – Local and global perturbations (shape parameter): This experiment follows the logic of the previous experiment with the difference that the parameter now shifting is the σ of the Log-Normal (i.e. the shape parameter). This shift is also 2 units. This perturbation introduces a much deeper disturbance to the distribution as Figure 6 shows.

The perturbation introduced has improved greatly the classifier's performance. Classifier robustness has also improved in global terms.

Behavior Prediction Through Automated Machine Learning

After a good understanding of the classifier's behavior gained in the previous experiments, now the authors apply AutoML to see how to further optimize these pipelines. This is a task where TPOT (Tree-based Pipeline Optimization Tool) can be of great assistance. TPOT is a Python library which tests different ML pipelines against a data set, see Randal (2016). Not only that, it also fine-tunes the hyperparameters of the prediction algorithms. TPOT is based on genetic programming and the user can configure TPOT to run based on a set of inputs. The reason behind the use of this strategy had to do with the very good performance that the Random Forest Classifier already exhibited in the previous experiments. Since TPOT is also tree based, this is a logic next step.

TPOT's Classifier uses randomness to find the best pipeline combination. Due to this, the same data set may be fed to TPOT using the same input parameters and produce different results. One way of making its results replicable, is to introduce the RANDOM_STATE variable in its input so that the same code may be ran twice and output the same pipeline as a result.

Although TPOT is immensely useful, it takes some tries to get it to run and finish the pipelines optimization. One of the causes for this behavior is that it requires a lot of computational resources and without the right configuration parameters it may crash. Another important aspect is that TPOT relies on other libraries such as Sci-Kit Learn, XGBoost or PyTorch which makes it particularly complex to set up. On the other hand, this is an advantage as these are popular ML libraries and their ML algorithms are open source and well-known within the scientific community.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. Data descriptive statistics for one of the stations of this line for the processing time feature for the data set of one shift

Statistics for Processing Time in the station	
Minimum	24.21 s
Average	48.49 s
Median	34.29 s
Maximum	164.25 s
Standard Deviation	29.40 s
Q1	29.61 s
Q2	34.23 s
Q3	50.67 s
Count	113

TPOT was applied for the same 5 data sets that were used in Experiment D. The results are presented in Table 9. This run was executed with GENERATIONS = 5, POPULATION_SIZE=10, N_JOBS = -1 and N_SPLITS = 10.

At a glance, it may look like there is some overfitting occurring. However, remember how in Experiment D already very good accuracies were achieved without the need to make any fine tuning to the hyperparameters of the Random Forest classifier. If the examples of Data Set 1 and Data Set 2 are to be taken into consideration, since they are optimized by the same pipeline, there is already an accuracy of 100%. This also describes the relative stability of the process itself that is being modeled.

In most scenarios, the pipeline is composed of an unique classifier. The exception is Data Set 4 where a stacking estimator is used. A stacking estimator uses different classifiers (in this scenario only two) and then uses the best outputs to maximize its accuracy. This is another advantage of using an AutoML tool such as TPOT. These combinations are done automatically. Otherwise, they might not occur to the Data Scientist or the number of combinations may be too cumbersome to test to a full extent.

Even though the authors have achieved some satisfying results in this section, there is now the need to test this methodology with real data in order to verify whether the classifier's characteristics hold.

THROUGHPUT PREDICTION ON MANUFACTURING PRODUCTION LINE

With the goal of testing the hypothesis at hand, in this work a data set which originates from an actual production line located in the Bosch Thermotechnology's Aveiro plant in Portugal was used. The data set was obtained from the Manufacturing Execution System (Bosch's Nexeed MES) and converted in order to match the structure of the general mathematical model. This line is comprised of a total of 18 workstations and it operates on 2 to 3 daily shifts. The yearly production for this line is around 50.000 parts. In Table 10 and Figure 7, the authors characterize one of these stations through some data statistics.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Histogram of the data characterized in Table 10

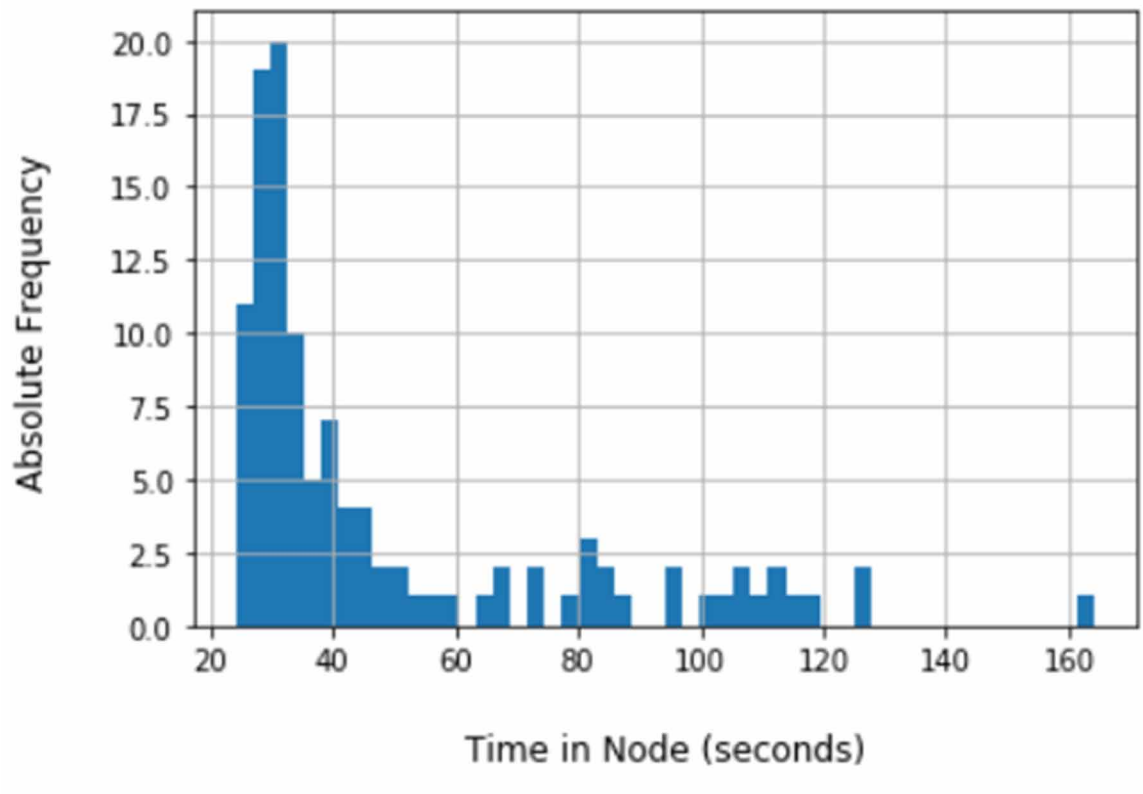
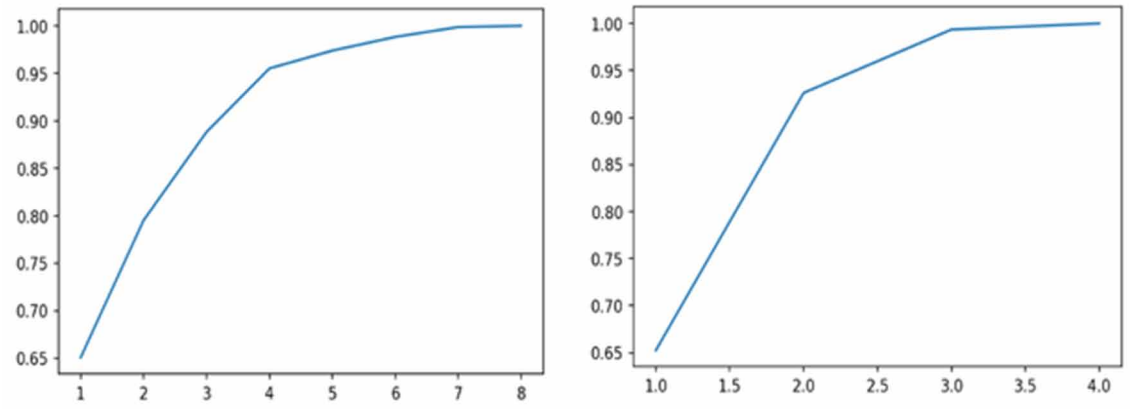
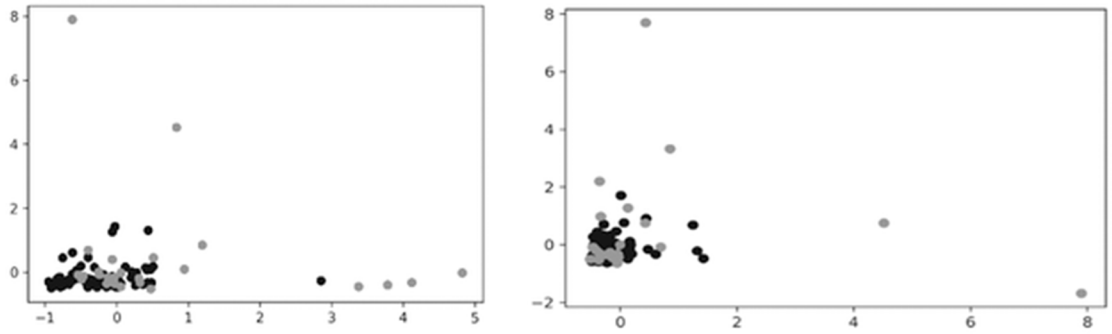


Figure 8 (left) and Figure 8 (right). Cumulative explained variance plotted (in the y-axis) and the number of features (in the x-axis) after PCA application using the Standard Scaler for the A and B data sets, respectively.



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9 (left, situation A) and Figure 9 (right, situation B). Darker values are below percentile 75%; lighter values are equal or above percentile 75%; the x-axis is related with the first principal component and the y-axis is related with the second principal component.



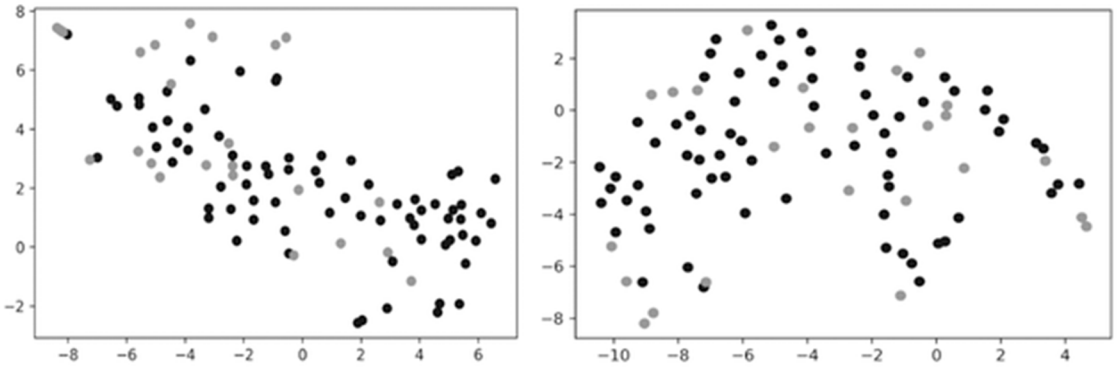
Dimension Reduction and Visualization

Dimension reduction is a standard approach to identify relevant features and remove others that do not give a significant contribution to the classifier's task. Although, there are many techniques for dimension reduction, the most popular technique is principal component analysis and its variants, see Liu (2007) and references within. Considering the mathematical characteristics of the data sets and the need to visualize the data topology (as reduction from the number of features to dimension 2), the t-distributed stochastic neighbor embedding was also implemented, see Maatens (2008).

Here, two data sets are considered: (A) having values for 8 elements; (B) having values for 4 elements; with the same accumulate production line transverse time with a total of 18 nodes.

PCA - Principal Component Analysis: PCA was introduced by Karl Pearson back in 1901 and there are many variants in the literature, e.g., robust PCA, kernel PCA, or localized PCA, see Ma (2019).

Figure 10 (left, situation A) and Figure 10 (right, situation B). Darker values below percentile 75%; lighter values equal or above percentile 75%; the x-axis is related with the x-coordinate of the net elements and the y-axis is related with the y-coordinate.



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11. Results with the line's data set where TPOT optimized different pipelines for each use case with a 75% percentile

	Accuracy	Pipeline
5 of 18 nodes	97.50%	StackingEstimator(estimator=MLPClassifier(alpha=0.01,learning_rate_init=1.0)), StackingEstimator(estimator=GaussianNB()), DecisionTreeClassifier(criterion="entropy", max_depth=9, min_samples_leaf=3, min_samples_split=11)
4 of 18 nodes	96.59%	StackingEstimator(estimator=DecisionTreeClassifier(criterion="entropy", max_depth=3, min_samples_leaf=20, min_samples_split=17)), BernoulliNB(alpha=10.0, fit_prior=True)
3 of 18 nodes	95.76%	DecisionTreeClassifier(criterion="entropy", max_depth=6, min_samples_leaf=10, min_samples_split=5)
2 of 18 nodes	86.36%	RandomForestClassifier(bootstrap=False, criterion="entropy", max_features=0.05, min_samples_leaf=1, min_samples_split=16, n_estimators=100)

Here, standard PCA is used. However, since PCA may be sensitive to feature scale and outliers, different scaling methods were tested. These feature transformations are comprehensively explained in SciKit-Learn's documentation, see Sci-Kit Learn (2020). It was observed that although the scaler choice is important in many scenarios, this was not the case here. In Figure 8, one can observe the results of such implementation for both data sets.

Figure 8 (left) shows that a relatively small number of elements is enough to characterize 90% of the data. Such conclusion is useful for applications in real situations.

Plots in Figures 9 are obtained by using PCA (with whiten and no renormalization for dimension reduction) for both situations. Although, a smaller number of elements characterize the data, the reduction to dimension 2 shows that the different classes are highly overlapped. Something expected since the elements nodes that are not in the data may be a source of increased processing time.

t-SNE – t-distributed Stochastic Neighbor Embedding: t-SNE is a very effective unsupervised nonlinear method for data visualization. While t-SNE preserves only small pairwise distances or local similarities, PCA is concerned with preserving large pairwise distances to maximize variance. Laurens illustrates the PCA and t-SNE approach pretty well using a Swiss Roll; if data has a spiral format distribution, PCA will connect points that are close in Euclidean distance sense but such points may be quite far when one follows the spiral structure.

From analyzing the results depicted in Figure 10, and since the unknown operation to predict has a linear structure, it is expected that data sets with a smaller number of elements will not lose too much information and by finding another well-suited ML pipeline, the classifier metrics will be similar to data sets with a higher number of elements information. Furthermore, there is no clear clustering structure for the label data (on the figures of both techniques) so linear ML techniques are not expected to perform well without applying an adapted lifting function to take advantage of the kernel trick. In other words, the classification problem is not trivial by itself.

Prediction Results

In Table 11 the results of predicting the correct label using from two to five sequential features are presented. This run was executed with GENERATIONS = 5, POPULATION_SIZE=10, N_JOBS = -1, N_SPLITS = 10 and N_REPEATS = 3, the same inputs as previously. In practice, this represents

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

how accurately one can predict the throughput time of a product based on the outcome at the early stages of the graph.

One conclusion seems to be that the simulator uses a probabilistic distribution which describes the physical phenomenon in a very good way. As it can be seen from Tables 10 and 11, the accuracy did not decrease significantly. This is a positive remark since drastic drops in accuracy are usually a common symptom of a rough theoretical approximation model to generate the synthetic data.

It might be surprising that the accuracy is so high in these scenarios. Since a label is being predicted which is tied with the throughput time of the graph, then the correlation that the ML algorithm is trying to find is actually a linear combination of the time spend in visible and hidden nodes/queues. Since TPOT explores a number of combinations between pipelines and the features provided, it is also not surprising that the results have such a high accuracy. However, a high accuracy may be the result of overfitting the training data with patterns specific to a particular (train, test) pair set.

From the start, the authors had always in mind that some challenges would probably arise, and this was indeed what was observed. Since the domain is organizational, there are several issues that need to be addressed.

Challenges for the Deployment of the Solution

With the increasing availability of data in many organizational departments, the urgency to put this data to good use is increasing at a steady pace. However, sometimes the challenges seem unsurmountable. From organizational to technical difficulties, there is still a long road ahead for most organizations to be able to get the most leverage out of their data and solutions as the one described with simulated data.

In this subsection, one will explore the types of challenges being faced. These challenges have been split into three categories: (i) organizational, (ii) development, and (iii) technical and data pipeline.

Organizational Challenges

Issues, controversies, and problems: As mentioned in Hartley (2019), digitalization is a process, and a long one indeed. From the start, it is important to convey to everyone in the organization the benefits and added value that these technologies bring. The human factor is generally one of the first barriers that digitalization faces.

People are at the center stage and they are the main drivers of this line of work. Data related techniques can be used in more traditional fields such as prediction or optimization, but they are starting to be increasingly used in other fields such as innovation, see Haefner (2021). As a result, most organizations are already including projects in their roadmaps which include techniques related with Machine Learning. However, putting these ideas into practice is very challenging in itself.

One of the first limitations encountered is related with talent management and enabling. It is often the case that very few people in the organization have the right know-how to conduct a long-term data strategy. With so many new technologies being released in very short timeframes, it is very easy to lose track or even know where to begin learning. This may occur even for highly trained and specialized engineers/IT personnel.

Moreover, even if the organization does have this kind of talent within its ranks, there is often a huge gap between collaborators which are experts in the organization's processes (being them from the manufacturing, logistics or financial areas) and the ones more focused on the technical side (infrastructure,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

mathematical techniques, etc.). This makes communication very hard and frustrating and causes many disturbances to the whole process.

Recommendations: The authors have explored previously how building know-how in technical and processual topics is a hindrance in today's organizations. There are different strategies to work around this challenge. One might be hiring someone with the right background in order to support the company to establish its own data strategy, not only in terms of the infrastructure but also for the overall, log-term data strategy. This could be a suitable strategy for Small and Medium Enterprises (SMEs). However, most often organizations will need a complete team to support this need. This sort of work involves almost the entire organization and requires extensive support from the higher management level in order to be able to work effectively due to the required changes it encompasses.

Another alternative, which is precisely the case at hand here, is the establishment of partnerships between local universities and organizations. This can translate into a tremendous advantage since both parties have a lot to contribute and benefit from working hand in hand. This combination of cutting-edge techniques and technologies with the field experience of collaborators of the organization is invaluable and the authors believe that this partnership will yield very significant results over time.

There was also the mention to the fact that there is often a huge gap between engineers and process experts. One recommendation to overcome such difficulties is to involve a third party which acts as a mediator between them. As data driven projects become more disseminated among organizations this need might tend to fade over time. At this point, it becomes clear that the organization is shifting towards more data awareness.

Development Challenges

Issues, controversies, and problems: Besides the challenges already mentioned which were more related with data acquisition and standardization, other important aspects also arise when it comes to developing and validating a ML framework which needs to process a great amount of information in real time. At the end of the first stage of development, the issue of verifying the correctness of the calculated metrics started to be a real issue. At first, the authors were using simulated data which was fed in real time to the system. Although this method does mimic the actual system in a comprehensive way, it was quite difficult to assess from the event's timestamps alone what was happening. Using real time data makes this process very time consuming and exceptionally prone to error. It is manageable with only one token in the system, but this is not a realistic approach to test and validate the framework's correctness.

Recommendations: In order to overcome this difficulty, a simulator was created to generate comma separated files which simulate the events that are expected to occur in the graph. This allows the authors to automatically associate to a certain token, the various timestamps and their calculated metrics. Another feature of this simulator is that it generates events following different probabilistic distributions whose parameters can be fully configured.

This simulator's use is actually two-fold: besides calculating the associated metrics, it also allows one to use the result to benchmark different ML pipelines against the data and to verify the accuracy of the prediction that can be achieved. When generating simulated data, multiple files are generated so that the ML pipelines can be tested against different data sets and thus assess the stability of the accuracy of the classifier.

Besides data validation, it is also worth mentioning that data privacy is a real concern and something which must be preserved at all times. For instance, in this scenario where a product's performance in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

a line is to be assessed, the identity of the operators working on the line is never to be disclosed and is completely unknown to the framework. An excellent reference on how ML models can have a hazardous influence in the day-to-day life of citizens and how to prevent these nefarious effects can be found in Kathy O'Neil's book, *Weapons of Math Destruction* (O'Neil, 2016).

Technical and Data Pipeline

Issues, controversies, and problems: There are no data driven projects without data. This is a fundamental truth, yet it is very easy to overlook whether a digital foundation exists in an organization. At its core, there are generally different Information Systems (ISs) which support all the business needs of these entities. Most often, these are Enterprise Resource Planning systems (ERPs) which combine a myriad of functionalities that all organizations need and rely on e.g. Human Resources or Financial processes. When one tends to look at the Manufacturing area, MESs are often a reality, but this might not be the case everywhere. This very diversified ecosystem means that having a data strategy is indeed mandatory. Otherwise, extracting actual information from such an infrastructure will be very time consuming and unproductive. This can be due to limited connectivity between the ISs and equipment yet one of the most common sources is the heterogeneity of the data sources themselves, which do not combine when one tries to fit them together.

Additionally, organizations usually face the ordeal of having to deal with legacy systems which have been in place for several years without a proper upgrading roadmap. Sometimes it may even be the case that the inner works of these systems are only known by a very limited amount of people which aggravates the problem. Most often, these are critical systems and the enterprise cannot fully function without them. This very entrenched dependency makes any effort of migration seem overwhelming due to the high risks it imposes. The same holds true for legacy equipment such as Programmable Logic Controllers (PLCs) and Industrial Personal Computers (IPCs) which have been installed many years ago and although these still supply critical functions to the manufacturing process, their connectivity is often quite limited thus not allowing these equipment to connect with the ISs in place. Most often, these devices need to be replaced by more recent ones with better functionalities, yet this transition is a very slow process. This happens not only due to technical challenges but also because of the disturbances its replacement may have in the productive process.

As one can picture and since the foundation is quite diverse, both in structure and even technologically wise, interfacing these systems with a ML framework is quite complex. Most organizations rely on flat file exports to fulfil these needs and having more advance features such as real time processing is not planned for the foreseeable future. Besides this, there is hardly ever an infrastructure in place which can meet the organization's data needs. This is often one of the first steps towards running data driven projects.

Finally, the fact that technologies are ever changing means that they become obsolete in a very short period of time. This is an added effort for organizations: besides needing to add more focus on maintaining their software up and running, technical skills are quickly deemed outdated as well. It was with this scenario in mind that a new approach was designed, one that could help organizations overcome some of these issues.

To sum up, the most prominent issues in this area are the lack of standardization in communication protocols and data structures, aged or missing technological infrastructure and poor interfaces between ISs.

Recommendations: The proposition is that a good data pipeline is not enough by itself, but together with a well characterized mathematical representation of the system relations and data, can solve multiple

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

data-driven problems across an organization. The main idea behind this framework is that by being highly concept oriented together with real time data, a generic tool is achieved which is useful for a myriad of different contexts and situations.

In order to fulfil the challenges previously mentioned, the authors have designed a framework which has the following characteristics:

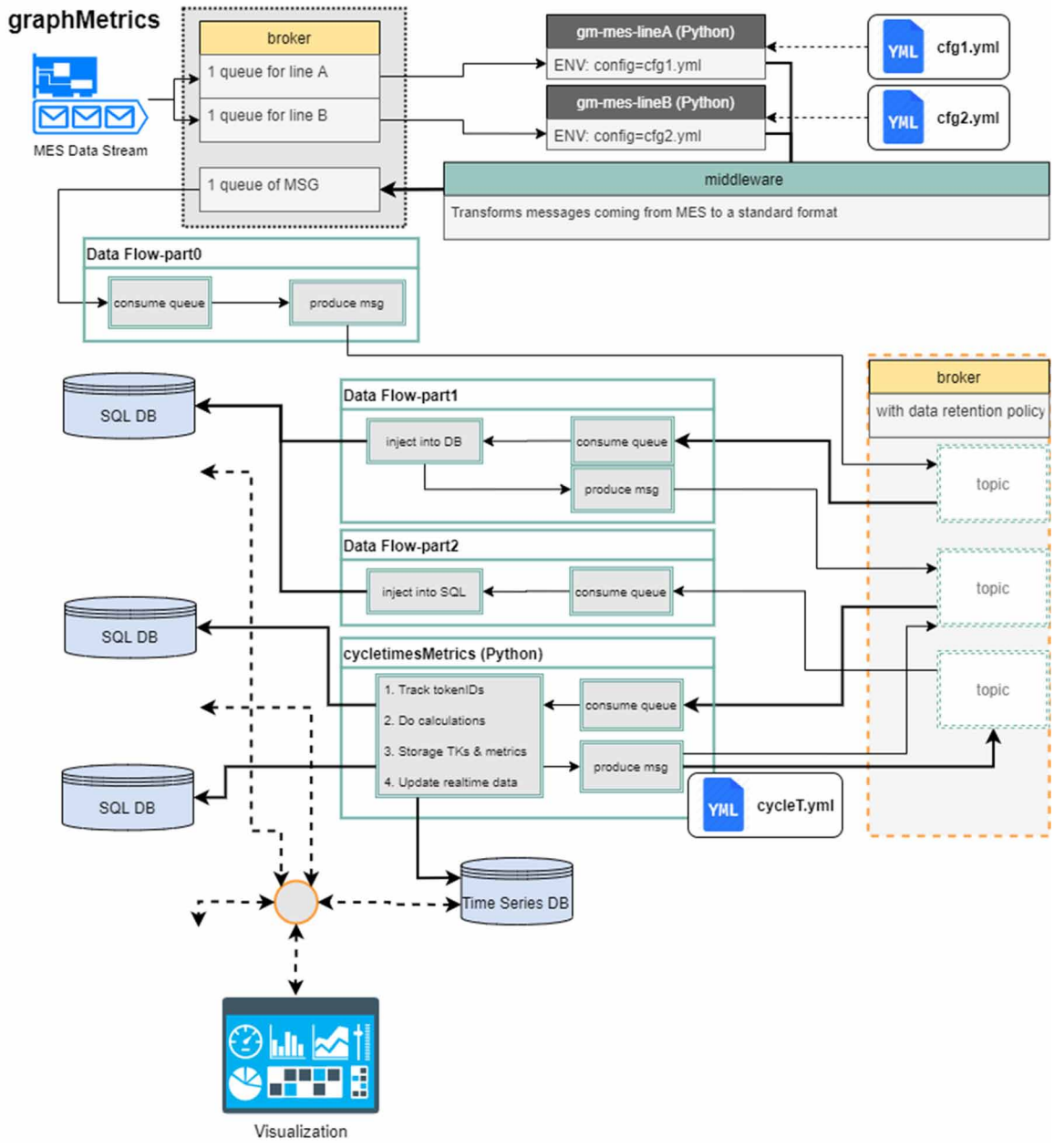
- Asynchronous Operation: Metrics for the graph are calculated asynchronously from the information currently reaching the system.
- High Configurability: The module relies on YAML files to allow the user to completely configure the process to be modeled by the abstract entity of the graph. For instance, if a different event defines the start of the graph, this can be configured.
- Easy Deployment: The architecture is implemented through the use of containers and the deployment process is carried out through a single text file. This makes deployment to production systems easy and consistent. This architecture enables a high degree of abstraction in relation to the Operation System (OS).
- Fault Tolerance: This was another important aspect taken into consideration and this is the reason why the usage of brokers and flow components is so extensive. In case of failure, no data gets lost. As soon as the service is available again, it is possible to resume operation starting from the moment of failure.
- Flexibility: This architecture was conceived with flexibility in mind so that not only systems such as MES can feed it but other data sources as well such as Internet of Things (IoT) devices. In order to provide this extended flexibility, a middleware was developed to allow for this uniformization.
- Machine Learning Pipeline Integration: This architecture enables us to integrate a standard ML pipeline, namely a training stage based on continuously ingested data and a prediction stage.
- Scalability: The system is container oriented and all data flows go through brokers so that it is horizontally scalable.

In Figure 11, the framework itself is pictured. Messages from the local MES are injected into a RabbitMQ service and a bridge microservice processes messages and injects them into a Kafka topic. These messages are processed by a middleware which filters out the messages which are not relevant and which translates the relevant ones into the standard data structure for the framework. After this pre-processing, the now standardized message is injected in a distinct Kafka topic where it will be processed by the Graph Metrics module. This module keeps track of all the tokens in the graph, their exact location and the metrics associated with them. In order to store all of this information, the authors rely on an ecosystem of databases for different use cases. For time series, a Prometheus variant was used. As an historical data repository, SQL DB was used but could be replaced by Apache Druid for OLAP. For visualization purposes, the recommended tool is Grafana yet there are more options available, such as Superset or Power BI.

Although this framework does solve the issue of data heterogeneity, it is recommended that organizations follow and implement a concrete roadmap for standardization and modernization of their ISs and manufacturing devices. Even though this is an enormous collective effort, it will certain pay-off in the long term since data heterogeneity involves added risk, higher maintenance costs and added development costs due to all the extra middleware components.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Proposed architecture for the general architecture for the graph based framework



CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Since the authors recognize the usefulness of the application of the techniques previously mentioned, they intend to make these functionalities readily and easily available to the end user. By disguising all of the inherent complexity, the plan is for the so-called process expert to be able to take the most out of the system in a proactive way. In order to achieve this, there is the need to come up with some creative interface methods such as asking the output results via email using natural language. Some Natural

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

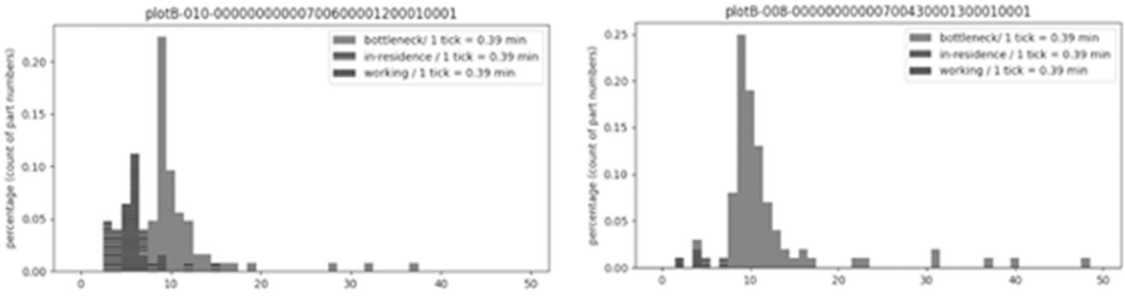
Language Processing methods would then translate the user's inputs into framework action triggers that produce and return reports answering users' questions.

Besides user interaction with the outputs from the system, there are three main other research directions.

One of these would be to use Priority Queues instead of FIFO (First In First Out) as a way of reducing the throughput time in a graph, in this particular case, applied to a PL. In this way, products which are known to have a behavior which is more variable would be processed in a separate way in order to prevent disruption to the graph's overall performance. This method would be first simulated and then applied to a real use case if the results are promising.

To understand the nature of the nodes durations, some additional figures are provided which show this distribution. As one can see from Figure 12 (left), there may be scenarios where the distribution does not really follow a log-normal and has a high number of outliers. However, these stations tend to be the exception rather than the norm. Most stations follow a pattern like the one seen on Figure 12 (right).

Figure 12 (left) and Figure 12 (right). On the left, plot of the duration of processing at Station 101 of the production line. On the right, plot of the duration of processing at Station 43 of the production line.



A second research direction would be to apply this same framework to bottleneck identification and optimization. Bottleneck detection is a methodology which could be easily applied to a framework like this since it takes first order metrics and creates a classification based on these metrics e.g. this station is a bottleneck or not in a given moment in time. The framework already provides the required metrics such as node and transition duration to do this analysis. This is a clear example of the practical usefulness of having a framework which has a stratified structure and where components are built on top of each other.

Finally, Root Cause Analysis for getting to the source of issues in local or overall graph performance can also be achieved through the use of such a framework. One could, for instance, understand why the duration in a certain station has a higher dispersion than all other stations. This could stem from various origins: the intrinsic nature of the process which takes place in a certain node, the interaction of the token with the node or the interaction of a third party with the node, for instance, the operator that may be working on a certain workstation.

In this chapter, the authors have explored a generic methodology to solve a wide range of problems which can be modeled by a graph. The starting hypothesis was whether a framework based on graph systems using AutoML could be efficient in solving problems of the nature already described. Based on the experiments that were run, the authors can conclude that this is a robust and accurate way of introducing ML to solve a panoply of real life problems. Using this methodology means that although there may be

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

a standard toolset, it does not mean, however, that all the problems are solved using the same classifier. The advantage of this methodology is precisely that there is no need to fix the classifier which is being used beforehand. That will depend on the intrinsic nature of the graph and processes which are being modelled. With that said, the pipeline is dynamic and can change over time as conditions also change. With this approach, the best of both worlds is combined to achieve both optimal flexibility and accuracy.

ACKNOWLEDGMENT

The authors acknowledge the support of Bosch Termotecnologia. The present study was also developed in the scope of the Project Augmented Humanity [POCI-01-0247-FEDER-046103], financed by Portugal 2020, under the Competitiveness and Internationalization Operational Program, the Lisbon Regional Operational Program, and by the European Regional Development Fund. The first author was supported by Bosch Termotecnologia. The second author was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA), through the Portuguese Foundation for Science and Technology, reference UIDB/04106/2020. The third author is funded by National Funds through the FCT - Foundation for Science and Technology, in the context of the project UIDB/00127/2020 (IEETA).

REFERENCES

Backus, P., Janakiram, M., Mowzoon, S., Runger, C., & Bhargava, A. (2006). Factory Cycle-Time Prediction with a Data-Mining Approach. *Semiconductor Manufacturing. IEEE Transactions on.*, *19*(2), 252–258. doi:10.1109/TSM.2006.873400

Buck, J., & Lee, E. (1993). Scheduling dynamic dataflow graphs with bounded memory using the token flow model. *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1, 429-432. 10.1109/ICASSP.1993.319147

Centomo, S., Panado, M., & Fummi, F. (2018). Cyber-Physical Systems Integration in a Production Line Simulator. *2018 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)*, 237-242. 10.1109/VLSI-SoC.2018.8644836

Chen, H., & Boning, D. (2017). Online and Incremental Machine Learning Approaches for IC Yield Improvement. *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. 10.1109/ICCAD.2017.8203857

Columbus, L. (2019). *10 Ways Machine Learning Is Revolutionizing Manufacturing in 2019*. https://www.forbes.com/sites/louiscolumbus/2019/08/11/10-ways-machine-learning-is-revolutionizing-manufacturing-in-2019/?sh=7e5167012b40

Domingos, P. (2017). *The Master Algorithm*. Penguin Books.

Eshuis, R., & Wieringa, R. (2003). Comparing Petri net and activity diagram variants for workflow modelling – a quest for reactive Petri nets. In H. Ehring, W. Reising, G. Rozenberg, & H. Weber (Eds.), Petri Net Technology for Communication-Based Systems, (vol. 2472). Springer.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Haefner, N., Wincent, J., & Gassman, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, 162.

Hartley, J., & Sawaya, W. (2019). Tortoise, not the hare: Digital transformation of supply chain business processes. *Business Horizons*, *62*(6), 707–715. doi:10.1016/j.bushor.2019.07.006

Lee, I., & Shin, Y. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, *63*(2), 157–170. doi:10.1016/j.bushor.2019.10.005

Liu, W., & Chang, C. (2007). Variants of Principal Components Analysis. *2007 IEEE International Geoscience and Remote Sensing Symposium*, 1083-1086, 10.1109/IGARSS.2007.4422989

Ma, J., Wang, A., Lin, F., Wesarg, S., & Erdt, M. (2019). A novel robust kernel principal component analysis for nonlinear statistical shape modeling from erroneous data. *Computerized Medical Imaging and Graphics*, *77*, 101638. doi:10.1016/j.compmedimag.2019.05.006 PMID:31550670

Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(86), 2579–2605.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.

Olson, Bartley, Urbanowicz, & Moore. (2016). Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, *GECCO '16* (pp. 485-492). Association for Computing Machinery.

Oosterbaan, R. J. (2019). Software for generalized and composite probability distributions. *International Journal of Mathematical and Computational Methods*, *4*, 1–9.

Russkikh, P., & Kapulin, D. (2020). Simulation modeling for optimal production planning using Tecnomatix software. *Journal of Physics: Conference Series*, *1661*, 012188. doi:10.1088/1742-6596/1661/1/012188

Sci-Kit Learn. (2020). *API Reference*. https://scikit-learn.org/stable/modules/classes.html#module-sklearn.preprocessing

Tušar, T., Gantar, K., Koblar, V., Ženko, B., & Filipič, B. (2017). A study of overfitting in optimization of a manufacturing quality control procedure. *Applied Soft Computing*, *59*, 77–87. doi:10.1016/j.asoc.2017.05.027

Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). A review of machine learning for the optimization of production processes. *International Journal of Advanced Manufacturing Technology*, *104*(5-8), 1889–1902. doi:10.100700170-019-03988-5

Wu, D., Jennings, C., Terpenny, J., Gao, R., & Kumara, S. (2017). A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. *Journal of Manufacturing Science and Engineering*, *139*(7), 071018. Advance online publication. doi:10.1115/1.4036350

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

AutoML: The process of automating the process of applying machine learning to real-world problems. AutoML covers the complete pipeline from the raw dataset to the deployable machine learning model.

Bottleneck: A bottleneck is a constraint in a manufacturing line where one single station is setting up the pace of the line. A bottleneck station is generally always working around the clock without any idle time. Bottlenecks are meant to be avoided since they hinder the performance of the overall line.

Graph: An abstract mathematical entity which possesses nodes and transitions. It is used to model systems with interactions between its various nodes. A manufacturing plant or a line are both examples of physical entities that can be described by a graph.

Manufacturing Execution System (MES): An Information System which is meant to control manufacturing operations in real time while also gathering data from the manufacturing process.

Token: An abstract mathematical entity which moves within a graph. It navigates between the graph's nodes.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 13
AI Meaning and Applications in the Consumer Sector of Retailing, Hospitality, and Tourism

Sandra Maria Correia Loureiro
https://orcid.org/0000-0001-8362-4430
ISCTE, BRU, Instituto Universitário de Lisboa, Portugal

Muhammad Ashfaq
https://orcid.org/0000-0001-7602-0647
Shenzhen University, China

Mariana Oliveira Berga Rodrigues
ISCTE, Instituto Universitário de Lisboa, Portugal

ABSTRACT

Artificial intelligence (AI) algorithms have been continuously adopted in businesses, such as retailing, hospitality, and tourism. The current chapter aims first to provide a conceptualization of evolution of AI. Second, it gives an overview of AI in business and finally points out examples of applications of AI in retailing, hospitality, and tourism. Third, the chapter offers suggestions for further research. This chapter is devoted to researchers and practitioners. The audience will benefit from an overview of AI meaning and potential research to be developed and also benefit from examples of other business that already employ such AI algorithms.

INTRODUCTION

Over the years, technology has been increasingly taking over our lives in a transparent but, yet tremendous way. Retail, hospitality, and tourism have started to adopt artificial intelligence (AI) to improve

DOI: 10.4018/978-1-7998-6985-6.ch013

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

customers satisfaction and enhance experiences with services and products. Additionally, the adoption of AI is also transforming the industries by considering the importance of productivity and how it can optimize processes and operations by making them more efficient and effective (e.g., Russell & Norvig, 2016; Loureiro, Guerreiro, & Tussyadiah, 2020; Romero, Ruiz-Equihua, Loureiro, & Casaló, 2021). However, there is still some hesitation regarding the future of this technology in the market and how to fully embrace it.

The number of companies and startups that are investing in AI keeps growing rapidly. It is a highly discussed topic both by managers and researchers. Nonetheless, AI and data will only have a significant and potential impact when its use is applied to the right case. The technological advancements have resulted in structural shifts and strategies within companies and businesses from the retail, hospitality, and tourism industries, that are increasingly operating more in a knowledge-based environment due to the information the technology is able to provide (e.g., Rust & Huang, 2014; Rust, 2020; Loureiro, Japutra, Molinillo, & Bilro, 2021). Through the application of software and algorithms, AI can interpret data and deliver more detailed knowledge regarding consumers' profiles, allowing managers to produce a more personalized and valuable experience. Based on the above considerations, the current chapter aims to (1) provide a conceptualization of evolution of AI, (2) give an overview of AI in business and finally points out examples of applications of AI in retailing, hospitality, and tourism and (3) provide suggestions for further research

BACKGROUND

Defining Artificial Intelligence

Since the term was first introduced in 1956, a precise definition of *artificial intelligence* (AI) has been a subject of debates and even causing a state of confusion, with researchers referring to AI from varying aims and approaches to AI research. Central to these has been the definition and meaning of *intelligence* (Legg & Hutter, 2007) and whether intelligence should be viewed as one ability or many (Kak, 2006). These issues are also rooted in the fact that things that are artificially intelligent differ from those that are naturally intelligent (Fetzer, 1990); artificial systems are significantly different to humans.

Definitions of AI also changed over time, especially as AI technologies develop rapidly recently. As observed by Russell (1997), some researchers in AI have focused on creating systems that emulate human cognition, while others on creating intelligence without concern for human characteristics. Still, many have endeavored to develop useful tools and artifacts without concern for intelligence. Russell and Norvig (2016) summarize the various definitions of artificially intelligent systems into four categories along two dimensions: reasoning–behavior dimension and human performance–rationality dimension. These are: (1) systems that think like humans, (2) systems that act like humans, (3) systems that think rationally, and (4) systems that act rationally. While tensions exist between approaches centered around humans and those around rationality, Russell and Norvig (2016) argue that each direction has yielded useful insights.

In an attempt to test universal intelligence, Legg and Hutter (2007, p. 12) suggest that "intelligence measures an agent's general ability to achieve goals in a wide range of environments." This implies three components of intelligence: an agent, environments, and goals. To achieve its goal, the agent will interact with the environments by sending and receiving signals to and from the environments (action

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

and perception). In terms of recognizing goals, they suggest that the environments can send signal that indicates how good the agent's current situation is (i.e., a reward). The agent's goal is thus to maximize the amount of reward it receives from the environments. Winston's (1992) earlier definition of AI as the "study of the computations that make it possible to perceive, reason, and act" (p. 5) seems to emulate this (acting rationally) approach. Similarly, Russell (1997) prefers to formally ground AI research based on bounded optimality as the desired property of AI systems. Bounded optimality refers to the capacity of machines to generate maximally successful behavior given the available information and computation resources. This approach takes into account the limitations of computational resources that are presumably responsible for deviation from human performance (Russell, 1997).

When AI programs need to interact with people, the issue of AI systems to act like humans becomes important. The Turing Test was created to test if AI system exhibits intelligent behavior, which is the ability to achieve human-level performance in all cognitive tasks. The Turing Test sets the foundation for various sub-fields of AI, as it requires AI systems to possess the following capabilities (Russell and Norvig 2016): *natural language processing* to communicate in a natural language, *knowledge representation* to store information, *automated reasoning* be used the stored information to answer questions and to draw new conclusions, and *machine learning* to adapt to new circumstances and to detect and extrapolate patterns. These capabilities are sufficient when physical interactions with human agents are absent. Additional capabilities are necessary to pass the Total Turing Test, which includes physical interactions: *computer vision* to perceive objects and *robotics* to move about.

According to Simmons and Chappell (1988), AI research in the mid-1950's focused on development of unique hardware (which was culminated in the perceptron) and development of programming tools for machines to emulate human behavior. The turning point in this phase was the project for automatic language translation, which unfortunately yielded unsatisfactory outcome. As a result, the field of AI was relatively dormant until the mid-1960s. This phase saw an achievement of impressive machine behavior, especially in the problem solving and diagnostic assistance (Simmons & Chappell, 1988). Projects such as Lisp Language (McCarthy, 1960) confirmed that given the right operating system and the right set of rules, machines could mimic human behavior rather accurately.

In the 1980's, the prototype of systems now called expert system (ES) was born from the work of researchers that used an architecture consisting of inference engines and knowledge base. This work pushes the application of AI technology to a wide range of practical problems. The key to AI systems acting humanly is the roles they assume relative to humans. By acting humanly, AI systems can be considered intelligent assistant to or replacement for humans. In business settings, the goals of AI research have been to create machines capable of completing low-level processing, speeding up a task while freeing human to do more creative work, and to create machines that augment human capability to improve human performance by acting as knowledge amplifiers (Simmons & Chappell, 1988).

Nowadays, business researchers conceptualize AI capabilities based on human skills it can replicate. For example, in the context of service, four different skills of AI were suggested: mechanical AI, analytical AI, intuitive AI, and empathetic AI (Huang & Rust, 2018). Mechanical AI is the skill associated with repetitive tasks, allowing machines to provide extreme consistency and reliability. Analytical AI is the skill associated with performing complex, yet systematic, consistent, and predictable tasks. This often involves processing a large amount of data and extracting patterns. Intuitive AI is the skill associated with performing complex, creative, holistic, and contextual tasks that need intuitive intelligence. Empathetic AI is the skill associated with providing psychological comfort for the wellbeing of humans.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Theoretical Models

e-Service through AI in the business field is becoming more and more prevalent in recent years. It has also become a hot topic for academic scholars. Scholars put considerable attention on AI across different fields. There are many technological models that have been used by researchers in their studies as the theoretical foundation in the context of AI. For example, when Ashfaq et al. (2020) studied users' satisfaction and their intentions to continue using AI-powered service agents (chatbots), they used the expectation-confirmation model (ECM), technology acceptance model (TAM), and information system success (ISS) model as the theoretical background. Sohn and Kwon (2020) also used the TAM to study users' behavioral intentions towards AI-based intelligent products. Similarly, drawing on the TAM, Belanche et al. (2019) investigated customers' adoption intention towards AI-based robo-advisors. Notably, the ECM was introduced by Bhattacherjee's (2001), while Davis (1989) proposed the TAM, and DeLone and McLean's (1992) suggested the ISS model. These models are the most commonly used to study users' adoption of technologies and decision making in technological and consumer behavior literature.

In addition to the theoretical models, scholars also used others relevant theories in the AI environment that include the expectations confirmation theory (ECT; Oliver, 1980), the unified theory of acceptance and use of technology (UTAUT; Venkatesh et al., 2003), the value-based adoption model (VAM; Kim et al., 2007), the theory of planned behavior (TPB; Ajzen, 1991), the technology-organizations-environment framework (TOE; Tornatzky and Fleischer, 1990), diffusion of innovation theory (DOI; Rogers, 1995), and Vroom's expectancy theory of motivation (Vroom, 1964). For instance, Brill et al. (2019) employed the ECT to study customer satisfaction with AI applications such as Siri and Alexa. Sohn and Kwon (2020) employed the UTAUT, VAM, and TPB to study users' intentions to use AI-based intelligent products. AlSheibani et al. (2018) used the TOE and DOI to investigate AI adoption at the firm level, while Chopra (2019) incorporated Vroom's expectancy theory to understand Indian shoppers' motivation to use AI. Figures 1-5 provide a graphical representation of some well-accepted models in the behavior literature.

EXAMPLES OF AI APPLIED IN INDUSTRY

Hospitality Industry

The hospitality industry is based on human services being, therefore, highly dependent on representation and customer reviews. Innovative technology can bring the customers' perspectives to the center of the operations (Loureiro et al., 2021). In fact, using AI methodologies can improve customers' experiences and simultaneously minimize the human workforce. Hence, over the recent years, hotels and other hospitality businesses have been investing more in technology in order to increase revenues and growth.

The application of AI in the hospitality industry holds the promise of transforming processes and services by making hotels more intelligent and including elements, such as, concierge robots, voice-activated services and automatic data processing (machine learning). These AI elements can consequently contribute to a better and more personalized service, which captures and provides information regarding customer purchases and profile (travel choices, travel patterns, preferences, payment methods, etc.).

Some of the first concierge AI robots applied to the hospitality industry were "Connie", launched by Hilton Worldwide and IBM, and "Mario" by Ghent Marriott Hotel. Both these robots were introduced in 2016, and developed to provide guests with restaurant recommendations, local tourist attractions and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. ECM (adapted from Bhattacherjee, 2001)

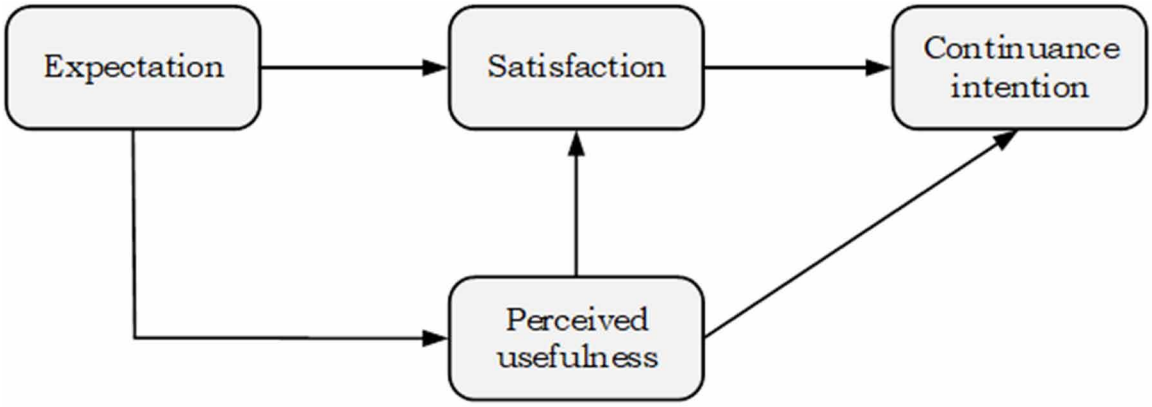


Figure 2. Tam (adapted from Davis, 1989)

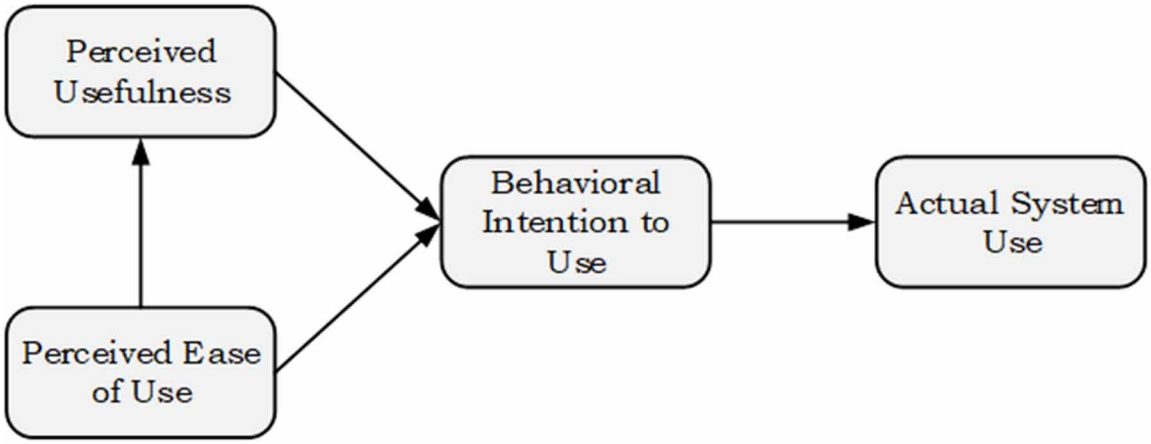
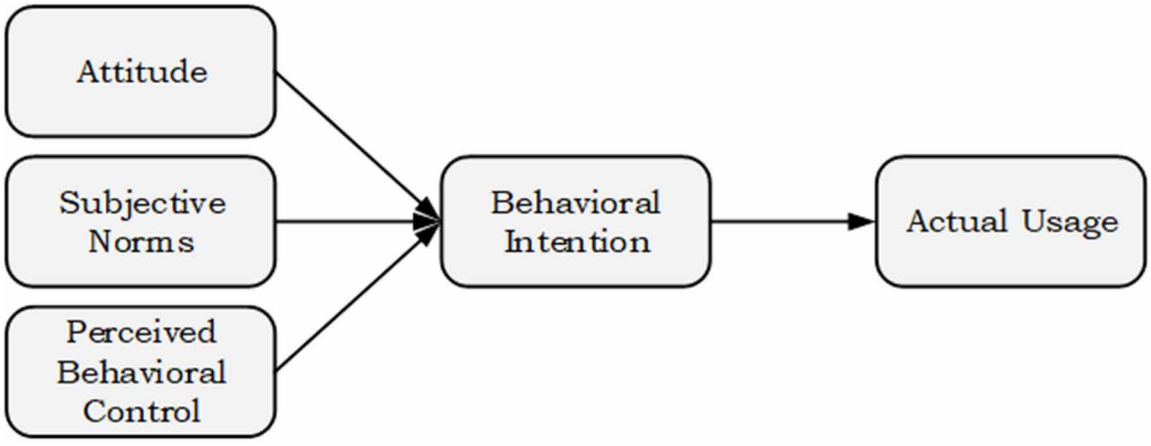
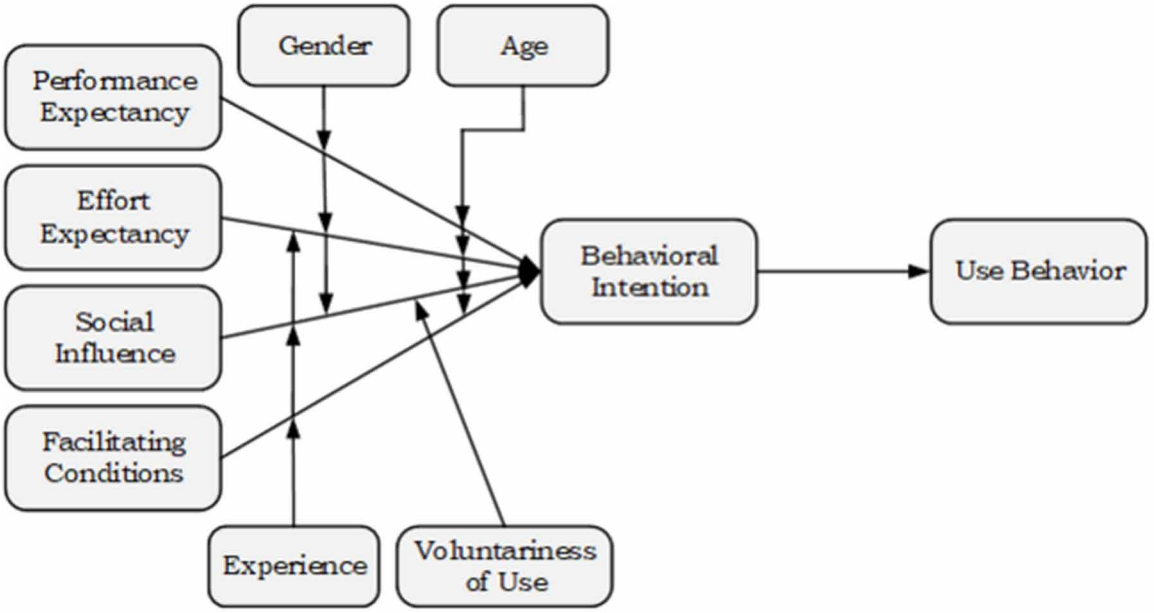


Figure 3. TPB (adapted from Ajzen, 1991)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

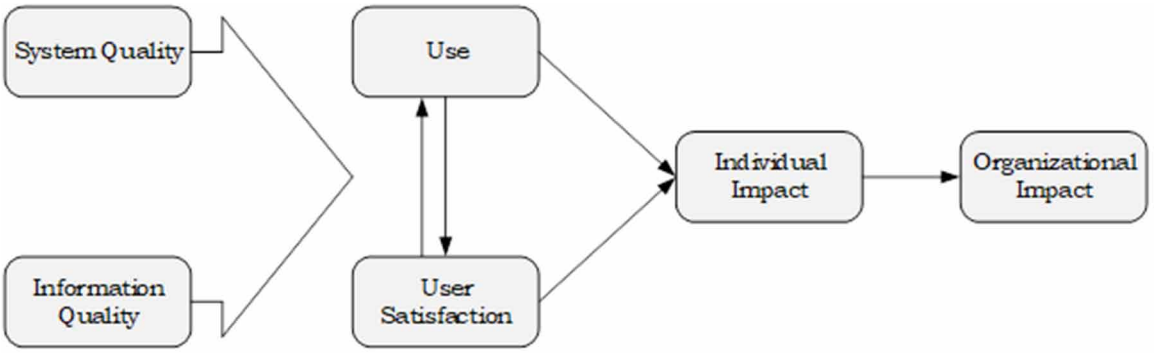
Figure 4. UTAUT (adapted from Venkatesh et al., 2003)



other hotel features. Hotels that are able to include AI concierge bots can offer a careful assistance and service at all time, thus, benefiting from customers' satisfaction. In Japan, Henna na Hotel is known for relying on robots as staff members to help and serve guests. Both the dinosaur robot and the humanoid female at the front desk are able to check guests in and out. Additionally, robotic porters are responsible for carrying the luggage, while the assistant robot – Churichan – is available in every room and can be used to provide information either about the room, hotel or even weather conditions.

In addition to AI robots, the hotel industry has been paying close attention to voice-activated AI assistance, by experiencing the technological abilities of machine learning software. This technology can use data analytics to learn about customers and, thus, help hotel managers and staff to create and deliver a more personalized experience and also to increase brand value. Clarion Hotel (Stockholm) and Wynn Las Vegas are some of the hotels known for integrating Amazon Echo, which provides an AI

Figure 5. 1SS Model (adapted from DeLone & McLean 1992)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

room assistant. Echo is easily activated by voice and ready to assure customers fulfill all their individual requirements throughout their stay, including control over the lights, room temperature, ordering room services, and providing information.

Hotels that invest in AI can not only benefit from more customer knowledge but are also able to design unique experiences that contribute differentiates them from competitors. For instance, Cosmopolitan of Las Vegas introduced an AI bot concierge called "Rose", in 2017. "Rose" has a good sense of humor and is flirty. In order for guests to interact with "Rose", they just need to text her number, which is given upon check-in. In addition to providing hotel services, "Rose" is also able to give restaurant and bar recommendations, a tour through the hotel's art collection, or deliver entertainment options and games, such as *Would You Rather* or *Two Truths and a Lie*.

Hotels are not the only ones investing rapidly in AI technology. In 2019, MSC Cruises presented Zoe, the first AI personal cruise assistant that is ready to promptly answer the most common questions during a cruise. Further, other cruises companies are also using AI to facilitate check-in (such as, Royal Caribbean's Intl.) and to customize the customers' onboard experience according to their digital identity (for instance, OceanMedallion on Princess Cruises).

Despite the many opportunities the application of AI presents in the hotel industry, it is still important to keep a critical eye regarding the technological gains. The hotel industry is highly characterized by the use of empathic intelligence, which is the type of intelligence that machines and technology still struggle to fully deliver. In fact, as research (Chan & Tung, 2019) contends, AI and respective robotic services do not necessarily result in a more positive brand experience. Even though technology may provide higher levels of sensory and intellectual experience, the same does not happen regarding the affective experience of the customers. Therefore, despite the technological changes we have been witnessing in the hotel industry, it is of high importance to consider how AI impacts operations and affective relationships with customers.

Tourism Industry

The tourism industry has also been increasingly applying Artificial Intelligence in order to improve service and the overall customer experience in tourism. Even though AI has an effective ability to enhance tourism experiences, similar to the hospitality industry, it does not dismiss the importance of human touch.

The emergence of AI in the industry has led to simpler and easy-to-follow travel arrangements, making services more automated and customized. The technology analyzes data and turns it into valuable information, allowing services to be more personalized, thus, improving the overall tourism experience.

The application of AI in the tourism industry can be a witness through different emerging technologies and benefit both customers and companies, agencies or organizations. For instance, chatbots have been used to quickly help customers by answering their travel questions. Expedia and Booking.com are some of the travel agencies that have included chatbots on their websites. From flights to places to visit and timetables, these intelligent chatbots are ready to guide customers and help them plan their time away from home. Further, since chatbots are able to respond to the most frequently asked questions, businesses using chatbots are able to optimize operations and services by focusing on other areas. Likewise, Skyscanner is also an online travel business using AI to enrich the customer experience. In order to facilitate booking processes, customers that use Alexa can easily access the cheapest flights to the desired destination and have it automatically booked. The technology allows for an easy and effective

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

communication between Alexa and Skyscanner, which is voice activated and makes processes simpler, thus, saving customers' time and struggles when searching for flights.

Additionally, the tourism industry is increasingly starting to rely more on machine learning to design consumer experiences based on personalized recommendations. In fact, in order to better understand the large amount of data the platform hosts, TripAdvisor has been using AI to analyze consumers' reviews on restaurants, hotel, activities and tourist sites. Consequently, according to those reviews, the platform is able to present more personalized recommendations with information that actually matters to each customer.

Further, tourist attractions, such as museums, are also trying to include more humanoid robots in order to enrich customers' experience. In fact, The Smithsonian Institute, a museum and research center, introduced Pepper, an AI robot, to engage with customers on a deeper level. Pepper is able to engage with visitors by telling stories, giving information on the museum and artwork, taking selfies, or even dancing. According to The Smithsonian Institute, since Pepper, the number of visitors in the museums increased and workers at the museum have more time to complete other tasks. Moreover, other museums (for instance, Anne Frank House, San Francisco Museum of Modern Art, Cooper-Hewitt Smithsonian Design Museum, and Akron Art Museum) have also been using chatbots to increase interactions, provide information and offer navigation assistance throughout the museum.

Retail Industry

The retail industry is characterized by the diversified operations it includes. Technological advancements and the fact that it is a highly competitive market leads to a need for constant change and adaption by the retailers in order to sustain advantages and meet the customers' expectations (Ashfaq, Yun,, Yu, & Loureiro, 2020).

The application of AI to the retail industry can transform and benefit a variety of aspects. When implementing AI, it is crucial to understand the results the technology might achieve not only at the present moment but also in the long run and how it can facilitate the constant need to adapt (Romero et al., 2021).

Artificial intelligence in retail can help customers discover new products and compare different products. LovetheSales.com is an online platform where customers can easily find and compare products from hundreds of retailers. In order to do so, the website uses Machine Learning technology, which enables the company to categorize, classify, and suggest products to customers. Moreover, AI can optimize operations and help retail managers make better and more profitable decisions according to data. Brands such as H&M and Morrison, rely on AI to make predictions regarding the stock needs, considering location, season, trends, etc. Similarly, Walmart is using autonomous scanners to keep up the inventory and analyze the stock on shelves. In addition, in order to innovate its drive-thru, McDonald's acquired an AI start-up that allows customers to have customized recommendations based on external factors, such as, traffic patterns, time of the day, and weather.

In addition to personalized recommendations, the implementation of AI technology can also have additional benefits. For example, Shop Runner relies on AI to help prevent fraud and provide customers with precise information that meets their interests. Furthermore, the use of AI in the retail industry can improve the customers' experience. Stich Fix, an online fashion website, uses algorithms to provide recommendations according to each person's style, budget, and size. Companies such as FarFetch, Amazon StyleSnap and Neiman Marcus, are using visual search. Meaning that customers are able to

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

upload pictures of clothes they like in real life and then the website generates results from the inventory, based on similar findings to those clothes. This allows customers to easily find the type of clothes they desire, without having to insert keywords that are too broad and may not lead to what they are looking for. Another brand example of how AI can be used to enhance the customers' experience is Nike. Nike launched a feature, called Nike Fit, which enables customers to try their shoe size and understand how it looks by combining AI with Augmented Reality (AR). This is a simple action that only requires the app and the phone's camera.

Regarding the in-store application of AI, Lowe's introduced an AI robot, LoweBot, which helps customers find the products they are looking for and can also be used to monitor the inventory. Further, Kroger presented Kroger EDGE, a technology that replaces traditional supermarket shelves with digital shelves.-

The application of AI in the retail industry can lead to several benefits that not only have a positive impact on the businesses but also contribute to a better and more personalized experience. The technology considers valuable data and customer insights, thus increasing sales and allowing customers to find the right products that meet their needs and desires. Overall, despite the high investment it requires, AI provides retailers with a more intelligent vision and strategy that makes processes more efficient and adaptable for the future.

FUTURE RESEARCH DIRECTIONS

While AI not only has certain benefits, it also has many deficiencies in these domains. Examples of such benefits include greater precision and accuracy, offering 27/4 customer services, convivence for both end-users and companies, significant error elimination, low cost, and a safer workplace or new possibilities for humans, such as space travel. Conversely, the decline in human ability due to over-use of AI, human autonomy, displacing human intelligence, and a higher unemployment rate may be indicators of disadvantages.

Considering these consequences of AI, academics scholars have suggested several research opportunities in the AI contexts that need to be explored in future research. For example, in the AI context, Huang and Rust (2018) reported four AI intelligences: mechanical, analytical, intuitive, and empathetic. They considered that (a) mechanical AI has the comparative advantage over humans, (b) analytical AI is "weak AI," (c) intuitive AI is "strong AI," and (d) empathetic AI is "experience" (e.g., the capability to experience things). According to Rust (2020), in mechanical tasks, AI is already very strong; however, it a long way to go with empathetic AI. Huang and Rust (2018) even stated that it is hard to programmed machines to experience emotions as persons do, despite the fact that the most advanced generation of AI is empathetic AI. Thus, future research is needed to understand how to program machines to experience feelings and emotions.

In addition to the future research trends and directions, Rust and Huang (2014) reported three main long-term trends: (1) a firm ability to communicate its customers, (2) a firm ability to collect accurate, precise, and detailed customers information and store it, and (3) a firm ability to the analyze the collected information regarding customers.

Furthermore, most of the previous studies on AI tend to adopt a very positive approach, either for organizations or users. However, AI-based automated technologies may offer limited customer experience and preference, as such technologies lack of human touch, unable to handle frustrated customers

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

or due to feelings of discomfort, privacy issues, perceived risk, job security concerns, uncertainty about performance, difficulty in interaction. Such technologies even have potential ethical and legal issues. For example, a driverless car with brake failure and pedestrians on the runway who would be killed if the car did not make any attempt to defend them might represent a moral dilemma. In this situation, guilt may be felt when the car killed someone, but it cannot be clear if it is the driver of the car, the company who made and built the car or if it is the car itself. The Law must also anticipate these circumstances and know their proper solution.

Tufekci (2016) points out that stereotypes have still been induced. Decision-makers say that we cannot avoid these difficult problems, but our activities should not be outsourced to robotics. Machine learning is mainly used by AI, and its decision-making is not fully transparent. No one will guarantee it will be processed and then understand how to learn information properly. Yet in this sense, future researchers must determine and use algorithms rather than abdicate and outsource. Moreover, society is concerned about the potential position of AI and how human beings can react to changes. It is clear that all of these reforms will benefit in terms of increasing productivity and livelihood, mitigating violence, and hazardous labor, etc. However, two main AI-related problems are usually listed, with the tremendous unemployment rate and the extinction of the human race some issues. The authors do, however, agree with the argument by Bostrom (2015) that we should incorporate human values intentionally and thoroughly into AI the same or even smarter than ourselves. In such a case, it may be better to adopt a broader point of view and discuss the trade-off effects associated with these new AI-based automated technologies. Thus, scholars should consider these concerns in future studies.

CONCLUSION

Traditionally, organizations have been used different ways to deliver excellent customer service to their clients, such as phone customer service, FAQ customer service, or email customer service. With the advancement of artificial intelligence, businesses are using AI-based customer service (such as chatbots). AI-based services bring new communication channels, and such channels offer new ways to interact with users to deliver good customer service around the clock. Thus, in this computer age, using technologies to provide uninterrupted services to the users at anytime from anywhere has always been a driving force for many players in the hospitality, tourism, and retail industries. Therefore, the implications of AI advancement should be regarded in the business field to deepen customer relationships. AI can play a significant role in consumers' experience of companies and building long-term relationships between companies and customers. The presence of AI is evident in today's business world, aiming to ease daily life, offering around the clock customer services, make individuals better off, enhance users' experience, and improve users' satisfaction level. AI is therefore currently used in and has an influence on many fields such as traveling, healthcare, education, society, public safety and security, employment and the workplace, or various home-use robots.

ACKNOWLEDGMENT

This work was supported by Fundação para a Ciência e a Tecnologia, grant UIDB/00315/2020.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. Advance online publication. doi:10.1016/0749-5978(91)90020-T

AlSheibani, S., Cheung, Y., & Messom, C. (2018). Artificial intelligence adoption: AI-readiness at firm-level. *Proceedings of the 22nd Pacific Asia Conference on Information Systems - Opportunities and Challenges for the Digitized Society: Are We Ready? PACIS 2018*.

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 1014734. doi:10.1016/j.tele.2020.101473

Belanche, D., Casaló, L. V., & Flavián, C. (2019). Artificial Intelligence in FinTech: Understanding robo-advisors adoption among customers. *Industrial Management & Data Systems*, *119*(7), 1411–1430. Advance online publication. doi:10.1108/IMDS-08-2018-0368

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly: Management*. *Information Systems*, *25*(3), 351. Advance online publication. doi:10.2307/3250921

Brill, T. M., Munoz, L., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*(15-16), 1401–1436. doi:10.1080/0267257X.2019.1687571

Chan, A. P. H., & Tung, V. W. S. (2019). Examining the effects of robotic service on brand experience: The moderating role of hotel segment. *Journal of Travel & Tourism Marketing*, *36*(4), 458–468. Advance online publication. doi:10.1080/10548408.2019.1568953

Chopra, K. (2019). Indian shopper motivation to use artificial intelligence: Generating Vroom's expectancy theory of motivation using grounded theory approach. *International Journal of Retail & Distribution Management*, *47*(3), 331–347. Advance online publication. doi:10.1108/IJRDM-11-2018-0251

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management*. *Information Systems*, *13*(3), 319. Advance online publication. doi:10.2307/249008

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95. Advance online publication. doi:10.1287/isre.3.1.60

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, *21*(2), 155–172. doi:10.1177/1094670517752459

Kak, S. (2006). Artificial and biological intelligence. *ACM Ubiquity*, *6*(42), 1–20.

Kim, H. W., Chan, H. C., & Gupta, S. (2007). Value-based Adoption of Mobile Internet: An empirical investigation. *Decision Support Systems*, *43*(1), 111–126. doi:10.1016/j.dss.2005.05.009

Legg, S., & Hutter, M. (2007). *Universal intelligence: a definition of machine intelligence*. Retrieved from https://arxiv.org/pdf/0712.3329.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2020). Artificial Intelligence in Business: State of the Art and Future Research Agenda. *Journal of Business Research*, *129*, 911–926. doi:10.1016/j.jbusres.2020.11.001

Loureiro, S. M. C., Japutra, A., Molinillo, S., & Bilro, R. G. (2021). Stand by me: Analyzing the tourist–intelligent voice assistant relationship quality. *International Journal of Contemporary Hospitality Management*. Advance online publication. doi:10.1108/IJCHM-09-2020-1032

McCarthy, J. (1960). Recursive Functions of Symbolic Expressions and their Computation by Machine, part I. *Communications of the ACM*, *3*(4), 184–195. doi:10.1145/367177.367199

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *JMR, Journal of Marketing Research*, *17*(4), 460–469. Advance online publication. doi:10.1177/002224378001700405

Rogers Everett, M. (1995). *Diffusion of innovations*.

Romero, J., Ruiz-Equihua, D., Loureiro, S. M. C., & Casaló, L. V. (2021). *Smart Speaker Recommendations: Impact of Gender Congruence and Amount of Information on Users' Engagement and Choice*. Frontier in Psychology., doi:10.3389/fpsyg.2021.659994

Russell, S. J. (1997). Rationality and intelligence. *Artificial Intelligence*, *94*(1-2), 57–77. doi:10.1016/S0004-3702(97)00026-X

Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson.

Rust, R. T. (2020). The future of marketing. *International Journal of Research in Marketing*, *37*(1), 15–26. doi:10.1016/j.ijresmar.2019.08.002

Rust, R. T., & Huang, M. H. (2014). The service revolution and the transformation of marketing science. *Marketing Science*, *33*(2), 206–221. doi:10.1287/mksc.2013.0836

Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics*, *47*, 101324. Advance online publication. doi:10.1016/j.tele.2019.101324

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). *Processes of Technological Innovation*. Lexington books.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *Management Information Systems Quarterly*, *36*(1), 157–178. doi:10.2307/41410412

Vroom, V. H. (1964). *Work and motivation*. Wiley.

Winston, P. H. (1992). *Artificial Intelligence* (3rd ed.). Addison-Wesley.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ADDITIONAL READING

Azarian, B. (2016), "A neuroscientist explains why artificially intelligent robots will never have consciousness like humans." Retrieved from https://www.rawstory.com/2016/03/a-neuroscientist-explains-why-artificially-intelligent-robots-will-never-have-consciousness-like-humans/

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, *63*(2), 215–226. doi:10.1016/j.bushor.2019.12.001

Bostrom, N. (2014). *Superintelligence: paths, dangers, strategies*. Oxford University Press.

Bostrom, N. (2015). What happens when our computers get smarter than we are? Retrieved from https://www.ted.com/talks/nick_bostrom_what_happens_when_our_computers_get_smarter_than_we_are

Colby, C. L., Mithas, S., & Parasuraman, A. (2016), "Service Robots: How Ready are Consumers to Adopt and What Drives Acceptance?" The 2016 Frontiers in Service Conference, Bergen, Norway.

Frankish, K. (2014). *The Cambridge Handbook of Artificial Intelligence*. CUP. doi:10.1017/CBO9781139046855

Goleman, D. (1996). *Emotional Intelligence: Why It Can Matter More than IQ*. Bloomsbury Publishing.

Gonzalez-Jimenez, H. (2018). Taking the fiction out of science fiction: (self-aware) robots and what they mean for society, retailers and marketers. *Futures*, *98*, 49–65. doi:10.1016/j.futures.2018.01.004

Ivanov, S., & Webster, C. (2019). Robots in tourism: A research agenda for tourism economics. *Tourism Economics*, *26*(7), 1065–1085. doi:10.1177/1354816619879583

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. doi:10.1016/j.artint.2018.07.007

Ng, I., & Wakenshaw, S. (2017). The internet-of-things: Review and research directions. *International Journal of Research in Marketing*, *34*(1), 3–21. doi:10.1016/j.ijresmar.2016.11.003

Pueyo, S. (2018). Growth, degrowth, and the challenge of artificial superintelligence. *Journal of Cleaner Production*, *197*(2), 1731–1736. doi:10.1016/j.jclepro.2016.12.138

Simmons, A. B., & Chappell, S. G. (1988). Artificial intelligence - definition and practice. *IEEE Journal of Oceanic Engineering*, *13*(2), 14–42. doi:10.1109/48.551

Sternberg, R. J. (1999). The Theory of Successful Intelligence. *Review of General Psychology*, *3*(4), 292–316. doi:10.1037/1089-2680.3.4.292

Sternberg, R. J. (2005). The Theory of Successful Intelligence. *Interamerican Journal of Psychology*, *39*(2), 189–202.

Tufekci, Z. (2016). Machine intelligence makes human morals more important. Retrieved from https://www.ted.com/talks/zeynep_tufekci_machine_intelligence_makes_human_morals_more_important

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative Intelligence: Humans and AI Are Joining Forces. *Harvard Business Review*, *96*(4), 114–123.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 14
Key Factors in the Process of Acceptance and Implementation of Artificial Intelligence in the Hotel Sector

Alfonso Infante-Moro
University of Huelva, Spain

Juan C. Infante-Moro
University of Huelva, Spain

Julia Gallardo-Pérez
University of Huelva, Spain

ABSTRACT

Artificial intelligence is one of the most innovative and trending technologies in the hotel sector. It is transforming the hotel sector into a novel environment in business-client relationships. For this reason, this study seeks to find, through a literary review and a causal study carried out by experts in technologies and hotels, using the methodology of fuzzy cognitive maps, what the determining factors in the decision of hotels to accept and implement artificial intelligence in their hotels and in their services are. The list obtained is made up of the following factors (in order of relevance in this decision): top management support, perceived reliability of the technology, security, relative advantage, support from information systems providers, technological organizational readiness, government pressure or incentives, compatibility, business partner pressure, customer pressure, complexity, pressure from competitors, perceived cost, characteristics of the leader or manager, and size of the company.

DOI: 10.4018/978-1-7998-6985-6.ch014

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Technologies have made organizations change the way they manage and provide services, so much so that their use in these functions within organizations has become almost essential. The same has also been observed in the hotel sector.

In the hotel sector, there are many technologies that have been successfully incorporated and have been a great revolution in the sector, but this study will focus only on one of them, Artificial Intelligence. One of the most innovative and trending technologies in this field, which is currently causing a great significant change in the hotel industry.

The term Artificial Intelligence was coined in 1956 by the computer scientist John McCarthy, who defined it as "the science and ingenuity of making intelligent machines" (Sosa Sierra, 2011), and today it is an interesting tool that is progressively transforming the hotel sector in a novel environment in company-client relations. This transformation focuses on digital interaction and greatly facilitates customer service with a great marketing strategy.

This application of Artificial Intelligence in the hotel sector aims to satisfy customer demand (the main axis of hotel management), especially with regard to the experience of guests during their stay and the automation of reservation processes. Among the services where this Artificial Intelligence can be observed and that could be carried out exclusively by this technology in the future are: check-in, check-out, reception, valet parking, chatbots or chat blogs for direct communication with the customer and service of rooms directly through the mobile device, among others.

Those hotels that seek to build an innovative and somewhat futuristic brand image, choose to implement the use of robots, improving efficiency, productivity and facilitating that staff can focus their attention on the customer experience. It does not mean that in the short term customers think of a robot as a bellboy or a receptionist, but the truth is that innovation and new technological projects help to improve performance, efficiency and productivity, and will make a difference to achieve success (Tecnohotel, 2018).

Applying a new resource to this environment can have both positive and negative points of view. Among the positives, the following stand out: benefits such as speed and precision, reductions in manual work, optimizations of personnel according to the workload, security improvements, energy savings through the monitoring of facilities, availability 24 hours a day and cost savings (for being cheaper). And among the negatives, the following stand out: lack of empathy and emotions essential to meet customer needs, technical failures that can lead to incorrect or misleading information, failures in data systems security, expensive and complex investment, and risks of requiring personnel for certain types of positions (Entorno turístico, 2020). Even so, it seems that the positive points of view are prevailing over the negative ones, since there are many hotels that are opting for its implantation.

For this reason, and in order to know which are the determining factors in the decision of hotels to accept and implement Artificial Intelligence in their hotels and in their services, this study is going to carry out a literary review that allows exposing a list of the factors that influence the adoption of Artificial Intelligence in the hotel sector and will develop a causal map of the system with them, in which experts in technologies and hotels will quantify and analyze the most influential factors when it comes to its acceptance by these hotels.

In the data capture (for the construction of the causal map) the interview will be used as an instrument, a causal map will be made for each of the interviewees and later a global causal map will be made with the average assessment of the experts, which will allow its analysis through the FCMappers tool,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

which will indicate the most influential factors, the factors that receive the most influence from the rest and the most relevant factors within the system.

In the following section, the importance of technologies in society and in the hotel sector is contextualized and the factors that influence when accepting an innovation or the use of new technological tools are detailed. We continue with the methodology used in this study and the analysis of the results, and conclude by pointing out the determining factors in the hotels decision to accept and implement Artificial Intelligence in their hotels and services.

BACKGROUND

Social changes and technological evolution have made information and communication technologies present in different sectors and areas of society (González-González et al., 2021; González-Zamar, Abad-Segura & Gallardo-Pérez, 2021; García-Peñalvo, 2021; García-Ruiz & Pérez Escoda, 2021; Viana & Peralta, 2021; Pangrazio & Sefton-Green, 2021; Martínez Pérez, Fernández Robles & Barroso Osuna, 2021; Pacheco-Cortés & Infante-Moro, 2020; Abad-Segura, González-Zamar, de la Rosa & Gallardo-Pérez, 2020; Flores-Fernandez & Martínez-López, 2020; Morales Salas, Infante-Moro & Gallardo-Pérez, 2020, 2019; González-Zamar, Abad-Segura & Belmonte-Ureña, 2020; Recio Muñoz, Silva Quiroz & Abricot Marchant, 2020). And in the business sphere, these technologies give companies a better competitive position, expanding their markets, optimizing their resources and adding value to their products and processes (Infante-Moro, Infante-Moro & Gallardo-Pérez, 2021b; Castrillón-Muñoz et al., 2019, 2020a, 2020b; Morales Salas & Rodríguez Pavón, 2020; Martínez-López, García-Ordaz, Arteaga-Sánchez & Infante-Moro, 2015).

Obviously, this use of technologies is also present in the hotel sector, where they are incorporated in the management, marketing and support of services that add value to customer experiences, this sector occupying the first positions in adoption, implementation and use of technologies (Domingo-Carrillo, González-Rodríguez & Chávez-Miranda, 2020; Infante-Moro, Infante-Moro & Gallardo-Pérez, 2020c, 2020d; Andrade & Fukuyama Sobata, 2020; Gonçalves et al., 2020).

Technologies are more than necessary in the hotel sector, a very competitive sector in which the vast majority of hotels make use of them (if not all) and make their use mandatory if they want to be competitive.

The use of technologies in management and marketing is part of this mandatory use that is no longer required to obtain a competitive advantage, but is required to coexist in the sector.

Now, if what is really sought is their use to achieve a competitive advantage, it should go to the implementation of technologies whose use is focused on supporting services that add value to customer experiences. These types of technologies that seek to personalize the experience of hotel clients (example: Artificial Intelligence and the Internet of Things) are a trend lately in the sector and not all hotels have them implemented (Europa Press, 2018).

Thus, in order to take measures to spread the use of these new technologies in the hotel sector and, in this case, spread the use of Artificial Intelligence in the hotel sector, this study focuses on locating which ones are the determining factors in the decision of the hotels to accept and implement Artificial Intelligence in their hotels and their services.

To do this, a literary review is first carried out to create a list of the most influential factors when accepting the use of new technological tools, collecting the factors that the authors Infante-Moro, Infante-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Moro and Gallardo-Pérez (2021a) pointed out in one of their studies on technological acceptance and that they distribute through the four contexts that make up the TOES framework (technological context, organizational context, environmental context and security context). The influence of these factors on technological acceptance is also confirmed by other studies, as can be seen in Table 1.

These factors represent the axis of distribution of the results of the study carried out to the experts in technologies and hotels below. This study tries to locate the determining factors in the decision of hotels to accept and implement Artificial Intelligence in their hotels and in their services.

METHODOLOGY

The Fuzzy Cognitive Maps (FCMs) methodology is one of the most relevant in the study of knowledge and, probably, one of the most used in recent times by researchers in their studies and projects (Papageorgiou & Salmerón, 2013). A tool that can be used in different situations or problems to identify, define and validate the factors of a system and identify the cause-effect relationships that exist between them, in order to propose strategies and help make decisions (Curia & Lavalle, 2011).

In the scientific field, it is used for its application: descriptive (to describe a system using the factors intervening in that system), explanatory (to explain the "why" of the behavior of the intervening factors in a system), reflective (to make decisions, as an instrument to support decision-making) and/or predictive (to predict the actions of the intervening factors in a system) (Codara, 1998).

For this reason, in this study, we decided to use this methodology, a methodology consisting of the visualization of a network of concepts and systemic factors linked by cause-effect relationships. This methodology allows to confirm or rule out the influences of the factors obtained in the literary review on the factors that influence when hotels decide to accept and implement Artificial Intelligence in their hotels and services. And, in addition, it allows to know whether or not there is influence between each of the factors that belong to said system and the degrees of all these influences (Maridueña, Leyva & Febles, 2016; Papageorgiou, Markinos & Gemtos, 2009).

And in order to obtain data on the degrees of influences existing between them, experts in technologies and hotels were selected as the "Population" under study (people who can get to know which are the most influential factors when hotels decide to accept and implement Artificial Intelligence in their hotels and services), and the sample size was set by the use of cumulative graphs of the number of new influencing factors with respect to the number of individuals belonging to the sampled interest group, as suggest U. Özesmy and S. L. Özesmy (2004) for this type of studies based on fuzzy cognitive maps (FCMs), which makes it possible to observe studies with a sample of 45 (Carley & Palmquist, 1992), 41 (Infante Moro, 2017), 30 (Infante-Moro, Infante-Moro & Gallardo-Pérez, 2021c), 29 (Carley & Palmquist, 1992), 8 (Infante-Moro, Infante-Moro & Gallardo-Pérez, 2020a, 2020b), 7 (Solana Gutierrez, Rincón Sanz, Alonso González & Garcia De Jalon Lastra, 2015) and up to 4 experts (Amat Abreu, Ortega Tenezaca & Yaguar Mariño, 2020).

In this study, the sample consisted of 30 experts and as can be seen in figure 1, this number was more than enough, all the participating experts selected these 15 factors as the variables of this system and none added new factors. Despite the fact that none added new influencing factors, it was decided to continue up to this number of interviewees so that the results were the most significant possible. And this significance also helped that there were not great differences between the values that the interviewees contributed for each relationship to be analyzed.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

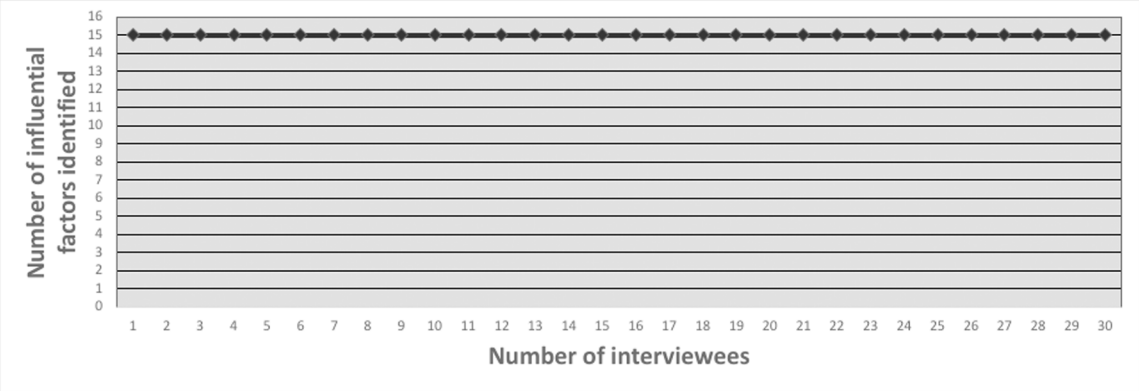
Table 1. Most influential factors when accepting the use of new technological tools (adapted from Infante-Moro, Infante-Moro and Gallardo-Pérez (2021a))

TECHNOLOGICAL CONTEXT	
FACTOR	**RELATIONSHIP**
The complexity	According to Sonnenwald, Maglaughlin and Whitton (2001), technological complexity is "the level of perceived difficulty of understanding and using technology". Berman, Kesterson-Townes, Marshall and Srivathsa (2012) state that adopted technologies must be user-friendly, easy to take advantage of and manageable to increase the opportunities for their adoption.
The compatibility	According to Rogers (2003), it can be defined as "the degree to which an innovation is perceived as consistent with current values, with the needs of potential adopters and their previous experiences".
The relative advantage	The relative advantage is "the degree to which an innovation is perceived as better than the idea it replaces" (Rogers, 2003). Therefore, the adoption of an innovation will be positively encouraged if its benefits are perceived to have advantages over existing systems and practices (Ekong, Ifinedo, Ayo & Ifinedo, 2012).
The perceived cost	The perceived cost refers to all types of expenses incurred by companies for the adoption of new technologies, such as hardware, software and systems integration (Premkumar, Ramamurthy & Crum, 1997).
ORGANIZATIONAL CONTEXT	
FACTOR	RELATIONSHIP
The characteristics of the leader or manager	Cerdán (2006) raises in his study three hypotheses related to the characteristics of the manager, which are the age of the manager, training, and degree of innovation; confirming the results that the age of the manager is among the main facilitators in the adoption of collaborative technologies, the university education of the manager among the neutral factors and the managerial attitude towards innovation at an intermediate level.
The perceived reliability of the technology	The reliability of a new technological tool is related to the trust of the company in new technologies (Tu, 2018).
The top management support	The degree of support from top management plays a very important role in the decision to adopt the new technology (Al-Shura, Zabadi, Abughazaleh & Alhadi, 2018). Previous studies such as that of Lin and Chen (2012) or that of Feuerlicht and Govardhan (2010) on the adoption of information technology innovation have shown that the support of top management is a critical factor in the decision of the organization for the adoption of a new technology, so that without this support organizations are more likely not to adopt new technologies.
The size of the company	The size of the company is another factor that can affect the adoption of information technology. Small companies and some medium-sized companies, despite the fact that they are more adaptable, are not willing to adopt new technologies (Lippert & Govindarajulu, 2006). In contrast, large companies have more opportunities than small and medium-sized companies (Al-Shura, Zabadi, Abughazaleh & Alhadi, 2018).
The technological organizational readiness	Organizational readiness is defined by Iacovou, Benbasat and Dexter (1995) as "the availability of the organizational resources necessary for adoption". Different researchers such as Mehrtens, Cragg and Mills (2001) or To and Ngai (2006) point out that the technological preparation of companies is important for the adoption of Information Systems (which include technologies) and covers not only the physical assets (technological infrastructure), but also specialized human resources.
ENVIRONMENTAL CONTEXT	
FACTOR	RELATIONSHIP
The pressure from competitors	Competitor pressure is "the level of pressure that the company may encounter from its competitors in the same area of the industry" (Laforet, 2011).
The business partner pressure	The decision of whether or not to adopt a new innovation in information technology is influenced by the history and past projects of a business partner (Gutierrez, Boukrami & Lumsden, 2015).
The customer pressure	Researchers like Kula and Tatoglu (2003) claim that companies innovate when their customers demand it.
The government pressure or incentives	Governments can play important roles in fostering the adoption of a new technological tool by using appropriate regulations or mandates as incentives to encourage participants to take action in the implementation of new technological tools (Tu, 2018).
The support from Information Systems (IS) providers	According to authors such as Rogers (2003) or Premkumar and Roberts (1999), the support of information systems providers refers to the support to implement and use information systems that a company obtains from external sources with technical experience.
SECURITY CONTEXT	
FACTOR	RELATIONSHIP
The security	Rolf H. Weber (2010) states that new technological tools have a significant impact on the security and privacy of the parties involved. Safety would be the feeling of the absence of danger or risk on the part of users when using a new technology (Weber, 2010).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

These experts in technologies and hotels were made up of two groups whose opinions had the same weight in the results: a group of 15 experts were university professors from different countries who have the title of doctor and among their lines of research are the technologies of the information and communications in hotels (these experts were selected to receive a scientific vision of the problem studied), and a group of 15 experts were responsible for the technology department in international hotels and hotel chains of 4- and 5-star hotels (these experts were selected to receive a professional vision of the problem studied).

Figure 1. Sample size selection process



The interview was the instrument used to obtain these data from the experts who formed the sample. This interview was conducted in order to know the opinion of the interviewees on the factors that influence when hotels decide to accept and implement Artificial Intelligence in their hotels and in their services, and on the degree of influence that these factors exert each other.

At the beginning of each of the interviews (these interviews were done individually and via video call), and as was done in the study by J. Solana (Solana Gutierrez, Rincón Sanz, Alonso González & Garcia De Jalon Lastra, 2015) in which fuzzy cognitive maps (FCMs) were used as a study technique, each interviewee received a template with the system factors, which served as an orientation guide for the interviewees (Table 2). In this way, it was intended to get the interviewees to take this list of factors and give the causal connections that existed between them. In addition to contributing, if they saw it appropriate, new factors.

But it is not the only table that was given to the interviewees at the beginning of each of the interviews, since not only was it necessary for this study to know whether or not these causal relationships existed, but it was also necessary to know the degree of the themselves. For this reason, they also received a semantic assessment table, whose semantic ranking of the connections is set according to the degree of causal influence that each of the interviewees think that one factor may have on another, and whether that influence is positive or negative (Mouratiadou & Moran, 2007) (Table 3). This table helped to quantify the degree of influences between the factors according to each of the interviewees.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Factors: Component concepts of the fuzzy cognitive map

TECHNOLOGICAL CONTEXT	
FACTOR	**DEFINITION**
The complexity	The level of perceived difficulty in understanding and using the Artificial Intelligence.
The compatibility	The degree to which the Artificial Intelligence is perceived as consistent with the values, needs, and prior experiences of potential adopters.
The relative advantage	The degree to which the Artificial Intelligence is perceived as better than the idea it replaces.
The perceived cost	The expenses incurred by the hotels for the adoption of the Artificial Intelligence.
ORGANIZATIONAL CONTEXT	
FACTOR	DEFINITION
The characteristics of the leader or manager	The age, training and degree of innovation of the manager.
The perceived reliability of the technology	The confidence of hotels in new technologies.
The top management support	The degree of support from the hotel management.
The size of the company	The size of the hotels.
The technological organizational readiness	The availability of the organizational resources necessary for adoption.
ENVIRONMENTAL CONTEXT	
FACTOR	DEFINITION
The pressure from competitors	The pressure for the installation in other hotels.
The business partner pressure	Investor pressure based on past technology adoption experiences.
The customer pressure	Pressure from customers.
The government pressure or incentives	The proper use of regulations or mandates as incentives to encourage participants to take action in the implementation of the Artificial Intelligence.
The support from Information Systems providers	Provider support to implement and use Artificial Intelligence.
SECURITY CONTEXT	
FACTOR	DEFINITION
The security	The feeling of the absence of danger or risk by users when using the Artificial Intelligence.

Once the data from the interviews were obtained, the existence (or not) of causal relationships between the proposed factors, an adjacency matrix was constructed per each interviewee that simulates the cause-effect unions between the factors formed by values of the interval [-1,1] (Banini & Bearman, 1998). And, subsequently, all these adjacent matrices or fuzzy cognitive maps (FCMs) were combined into a single matrix or collective map, in which the average value of the individual matrices was assigned to each factor. Through this, the degree of influence of these factors in the hotels' decision to accept and implement Artificial Intelligence in their hotels and services was obtained. The system factors were classified according to their influences on the rest of the factors and the outdegree, indegree and centrality indicators were obtained using FCMappers software (Bachhofer & Wildenberg, 2010).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Semantic assessment table: interpretation of cause–effect relationships

Value	Semantic Relation
1	Very strong positive
0.9	
0.8	Strong positive
0.7	
0.6	Medium positive
0.5	
0.4	Weak positive
0.3	
0.2	Very weak positive
0.1	
0	There is no relationship
−0.1	
−0.2	Very weak negative
−0.3	
−0.4	Weak negative
−0.5	
−0.6	Medium negative
−0.7	
−0.8	Strong negative
−0.9	
−1	Very strong negative

RESULTS

The collective cognitive matrix or collective map based on the existing influences between the factors of this system was made up of 15 factors and 115 causal connections resulting from 30 matrices corresponding to 30 interviewed experts (Table 4).

Thirteen of these factors exert and are influenced by the other factors, so these variables are ordinary variables, and two of these factors (the size of the company and the characteristics of the leader or manager) only exert influence on the rest of the factors but are not influenced, so these variables are transmitter-type variables.

When processing these connections with the FCMappers software, the structure of this fuzzy cognitive map was analyzed through the following indicators (Özesmi & Özesmi, 2003):

- The ‘outdegree’ indicator indicates the degree of influence of a factor on the rest of the factors. A very influential factor presents a high outdegree.
- The ‘indegree’ indicator indicates the degree to which a factor is influenced by the rest of the factors. A highly influenced factor presents a high indegree.

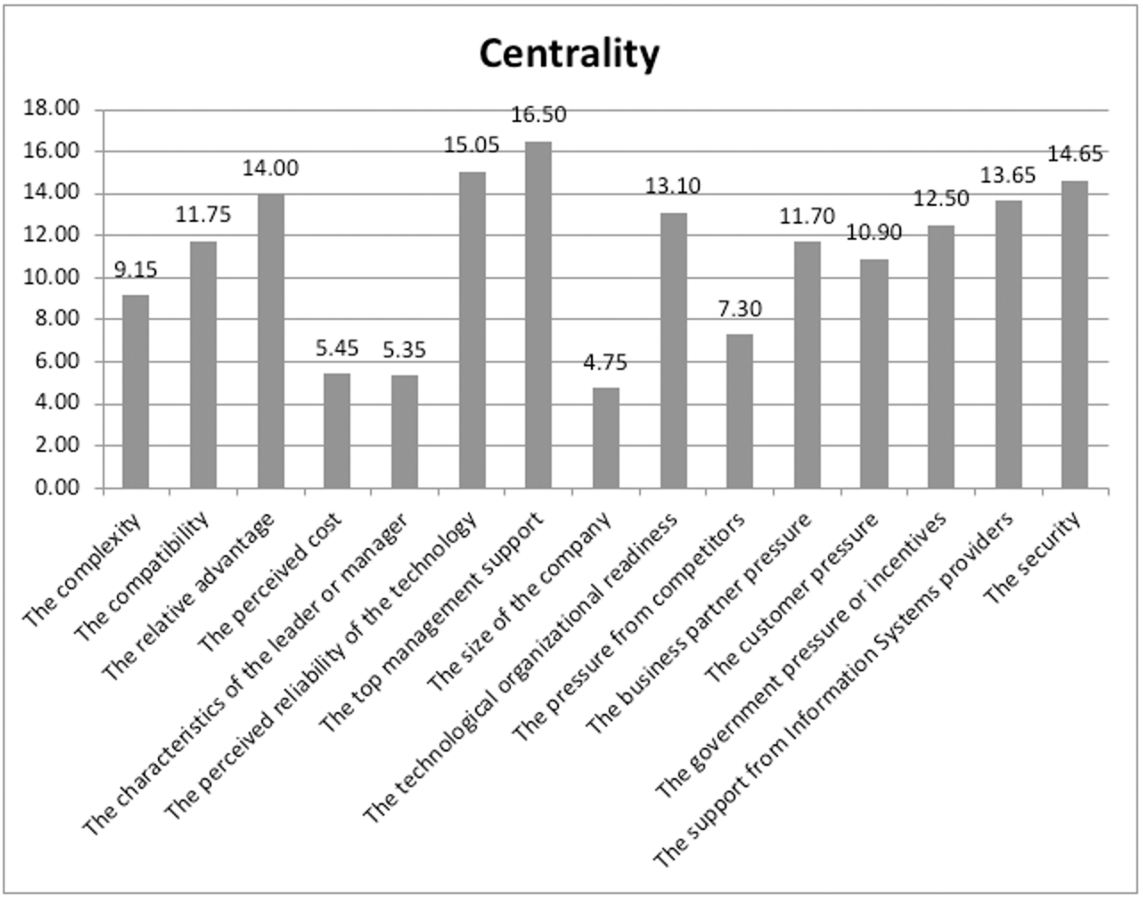
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Adjacent matrix of the collective fuzzy cognitive map

	The complexity	The compatibility	The relative advantage	The perceived cost	The characteristics of the leader or manager	The perceived reliability of the technology	The top management support	The size of the company	The technological organizational readiness	The pressure from competitors	The business partner pressure	The customer pressure	The government pressure or incentives	The support from Information Systems providers	The security
The complexity	0.00	-0.40	-0.70	0.00	0.00	0.70	-0.75	0.00	0.00	0.00	0.00	0.00	0.70	1.00	-0.70
The compatibility	-0.40	0.00	0.80	0.00	0.00	0.60	0.90	0.00	0.00	0.60	0.60	0.00	0.00	0.00	0.00
The relative advantage	0.00	0.70	0.00	0.00	0.00	0.60	0.95	0.00	0.00	0.65	0.80	0.70	0.00	0.00	0.00
The perceived cost	0.00	0.00	-0.85	0.00	0.00	0.00	-0.90	0.00	0.00	0.00	0.70	0.00	0.00	0.00	0.00
The characteristics of the leader or manager	0.00	0.30	0.60	0.00	0.00	0.70	0.95	0.00	0.70	0.00	0.00	0.00	0.80	0.80	0.50
The perceived reliability of the technology	-0.60	0.70	0.80	0.00	0.00	0.00	0.90	0.00	0.85	0.00	-0.70	0.00	0.00	-0.20	1.00
The top management support	0.00	0.75	0.40	0.00	0.00	0.90	0.00	0.00	0.90	0.00	-0.40	0.00	0.40	0.60	0.70
The size of the company	0.00	0.40	0.65	0.80	0.00	0.00	0.00	0.00	0.80	0.00	0.00	0.00	0.70	0.70	-0.70
The technological organizational readiness	-0.80	0.80	0.75	-0.60	0.00	0.85	0.80	0.00	0.00	0.00	-0.40	0.00	0.40	-0.60	0.75
The pressure from competitors	0.00	0.70	0.80	0.00	0.00	0.75	0.90	0.00	0.70	0.00	0.80	0.80	0.00	0.00	0.00
The business partner pressure	0.00	0.60	0.70	0.00	0.00	0.60	0.90	0.00	0.80	0.00	0.00	0.00	0.40	0.40	0.40
The customer pressure	0.00	0.90	0.95	0.00	0.00	0.80	0.90	0.00	0.60	0.60	0.80	0.00	0.70	0.65	0.90
The government pressure or incentives	0.70	0.00	0.00	0.70	0.00	0.90	0.80	0.00	0.60	0.00	-0.40	-0.70	0.00	0.90	1.00
The support from Information Systems providers	-0.90	0.80	0.70	0.90	0.00	0.90	0.90	0.00	0.40	0.00	-0.40	0.00	0.90	0.00	1.00
The security	-0.80	0.80	0.90	0.00	0.00	1.00	0.90	0.00	0.00	0.00	-0.90	-0.90	0.80	0.00	0.00

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Indicators of the fuzzy cognitive map: Centrality



- And the 'centrality' indicator indicates the degree of participation or importance of a factor in the system, it is the sum of the outdegree and indegree indicators.

In this way, the most important factors in this decision are those that have higher centrality values, since they regulate the dynamics of the system (Özesmi & Özesmi, 2004) (Figure 2).

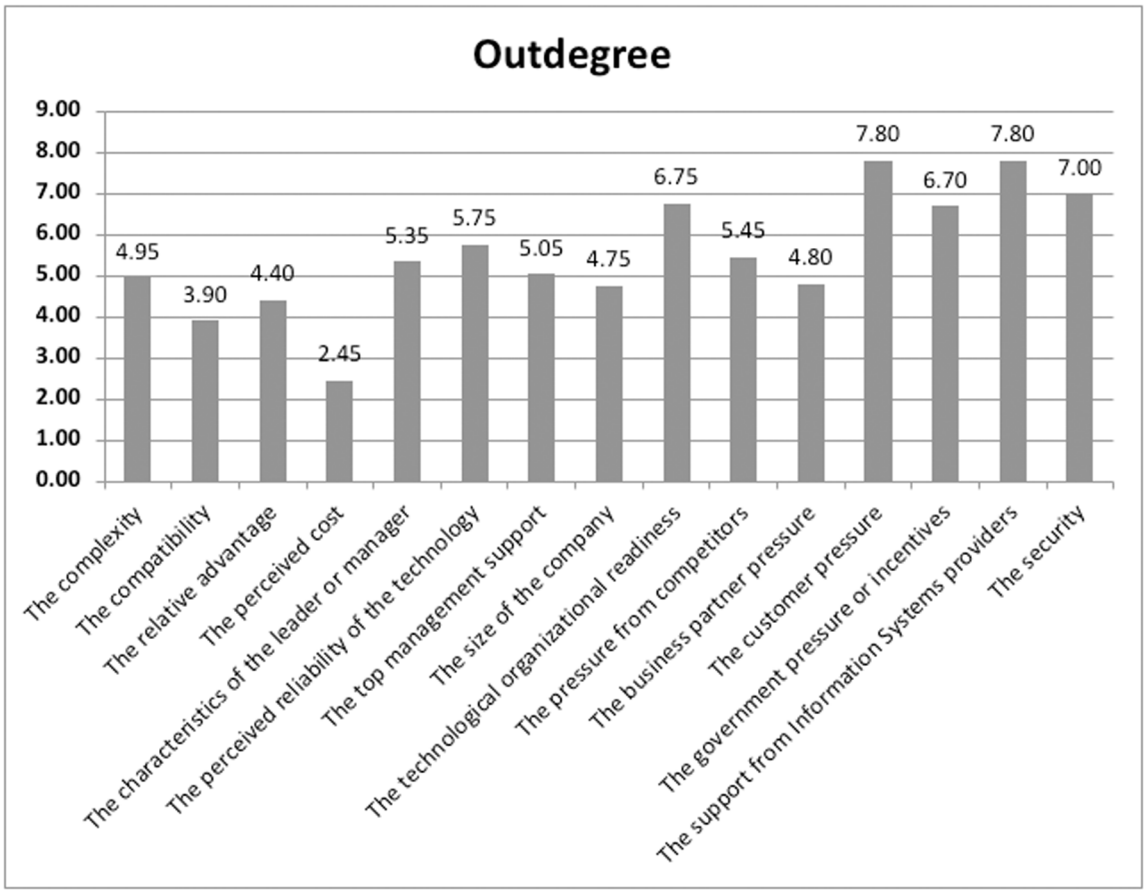
And between outdegree and indegree: the factors that have high outdegree values strongly influence the rest of the factors (customer pressure, support from Information Systems providers and security), and the factors that have high values of indegree are strongly influenced and dependent on the rest of the factors (top management support, relative advantage, and perceived reliability of the technology) (Figures 3 and 4).

With all this, the variables or factors with the greatest centrality, outdegree and indegree are shown in the following table (Table 5).

The factor with the greatest centrality is the top management support, which is the ninth most influential in the rest of the factors (outdegree = 5.05) and the most dependent on the other factors (indegree = 11.45). And the factor with the least centrality is the size of the company, which is the fourth with the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Indicators of the fuzzy cognitive map: Outdegree



lowest degree of outdegree and the one with the lowest degree of indegree. The main factors ordered according to the centrality indicator are shown in Figure 2.

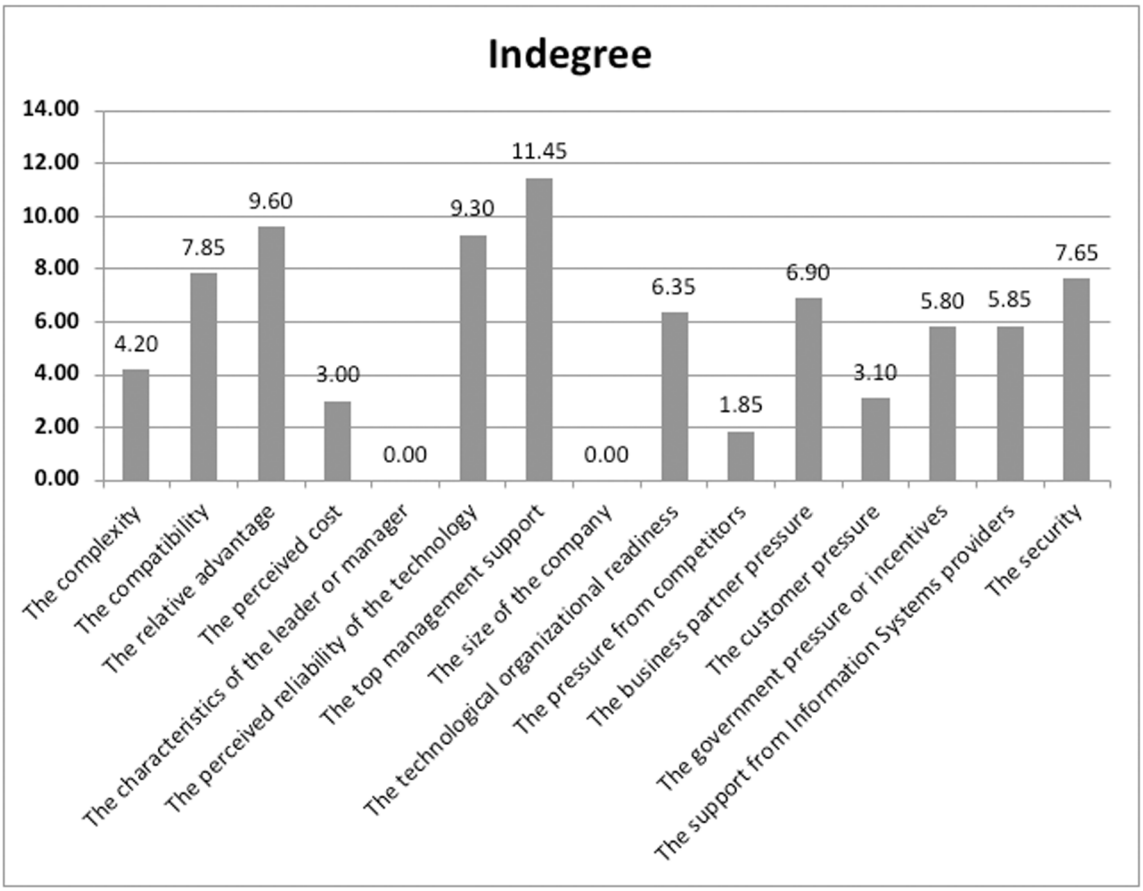
The factors that exert the most influence on the rest of the factors are: customer pressure, support of Information Systems providers and security. And the most dependent factors are: top management support, relative advantage, and perceived reliability of the technology. The main factors ordered according to the outdegree and indegree indicators are shown in Figures 3 and 4.

FUTURE RESEARCH DIRECTIONS

This research is limited to the study of the factors that can be collected in the contexts of the TOES framework (technological context, organizational context, environmental context and security context), but this limitation can be eliminated in future studies and extend this study to other types of factors that also influence this decision, such as the study of the motivational factors that can influence this decision or the study of the strengths and weaknesses that the different applications of said technology may have.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Indicators of the fuzzy cognitive map: Indegree



CONCLUSION

More and more hotels are choosing to implement Artificial Intelligence in their hotels and services. For this reason, this study has tried to locate the determining factors in the decision of hotels to accept and implement Artificial Intelligence in their hotels and in their services, allowing to expose a list of influencing factors when accepting the use of new technological tools (in this case, when hotels accept this technology as a complementary tool) and determine which are the most influential or decisive when it comes to being accepted by these hotels.

Table 5. Factors with higher indicators of centrality, outdegree and indegree

Centrality	Outdegree	Indegree
The top management support	The customer pressure	The top management support
The perceived reliability of the technology	The support of Information Systems providers	The relative advantage
The security	The security	The perceived reliability of the technology

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The list consisted of the following factors:

- In the technological context: complexity, compatibility, relative advantage and perceived cost.
- In the organizational context: characteristics of the leader or manager, perceived reliability of the technology, top management support, size of the company and technological organizational readiness.
- In the environmental context: pressure from competitors, business partner pressure, customer pressure, government pressure or incentives, and support from Information Systems providers.
- And in the security context: security.

And the most decisive factors when hotels decide to accept and implement Artificial Intelligence in their hotels and services are (in order of relevance): top management support, perceived reliability of the technology, security, relative advantage, support from Information Systems providers, technological organizational readiness, government pressure or incentives, compatibility, business partner pressure, customer pressure, complexity, pressure from competitors, perceived cost, characteristics of the leader or manager, and size of the company.

With what in general lines, it can be said that a system is observed in which the two factors that exert the most influence on the rest of the factors are the customers pressure (something normal in the service sector) and the support from Information Systems providers (due to the complexity of the technology), followed by other factors again related to the technology to be implemented and its complexity (these factors are: security, technological organizational readiness and government pressure or incentives). And a system in which the top management support, the relative advantage and the perceived reliability of the technology are dependent on the rest of the factors, and become three of the four most important factors of this system or decision making (together security), placing the value of this decision on the support of the hotel management, the trust of the hotels in new technologies, the feeling of absence of danger or risk on the part of the users when using Artificial Intelligence and the plus that the application of this technology would give to the idea that it replaces. To the detriment of another factor such as the perceived cost, which is the one that has the least influence on the rest of the factors and one of the three that least influences this decision (together with the characteristics of the leader or manager, and the size of the company), since hotels are willing to bear the cost of its implementation if its advantages and benefits are considerable.

The latter is something that does not coincide with the study by Tu (2018), where the perceived cost is considered one of the most influential factors when implementing a new technological tool in companies in general, since they consider that the costs could become significant barriers to adoption. Something that is omitted in the hotel sector, due to what has been said above.

And regarding the three most determining factors in this decision, it does coincide with what can be seen in other sectors, where the top management support, the perceived reliability of the technology and security, are considered among the factors with the most influence in this decision (Al-Shura, Zabadi, Abughazaleh & Alhadi, 2018; Tu, 2018; Lin & Chen, 2012; Feuerlicht & Govardhan, 2010; Weber 2010).

With this, it has been tried to glimpse what are the determining factors in the decision of hotels to accept and implement Artificial Intelligence in their hotels and in their services, which has shown a process in which the top management support is the factor more decisive. This factor depends on almost all factors, some are taken for granted in the sector and management already knows their advantages, but others can influence when increasing the support of management (those related to the technical aspects

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the technology to be implemented and its complexity, something that a good disposition of information about the technology to be implemented and good support from suppliers can help).

REFERENCES

Abad-Segura, E., González-Zamar, M. D., de la Rosa, A. L., & Gallardo-Pérez, J. (2020). Management of the digital economy in higher education: Trends and future perspectives [Gestión de la economía digital en la educación superior: tendencias y perspectivas futuras]. *Campus Virtuales*, *9*(1), 57–68.

Al-Shura, M. S., Zabadi, A. M., Abughazaleh, M., & Alhadi, M. A. (2018). Critical success factors for adopting cloud computing in the pharmaceutical manufacturing companies. *Management and Economics Review*, *3*(2), 123–137. doi:10.24818/mer/2018.12-01

Amat Abreu, M., Ortega Tenezaca, D. B., & Yaguar Mariño, J. J. (2020). Determination of the degree of influence of the climatic factors of vulnerability of the agricultural sector with neutrosophic techniques [Determinación del grado de influencia de los factores climáticos de vulnerabilidad del sector agropecuario con técnicas neutrosóficas]. *Investigação Operacional*, *41*(5), 699–705.

Andrade, J., & Fukuyama Sobata, M. (2020). Content aspects valued by users of travel blogs. Insights from brazilian travelers. *International Journal of Information Systems and Tourism*, *5*(1), 7–17.

Bachhofer, M., & Wildenberg, M. (2010). *FCMappers*. http://www.fcmappers.net

Banini, G. A., & Bearman, R. A. (1998). Application of fuzzy cognitive maps to factors affecting slurry rheology.[. *International Journal of Mineral Processing*, *52*(4), 233–244. doi:10.1016/S0301-7516(97)00071-9

Berman, S. J., Kesterson-Townes, L., Marshall, A., & Srivathsa, R. (2012). How cloud computing enables process and business model innovation. *Strategy and Leadership*, *40*(4), 27–35. doi:10.1108/10878571211242920

Carley, K., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models.[. *Social Forces*, *70*(3), 601–636. doi:10.2307/2579746

Castrillón-Muñoz, A. J., Infante-Moro, A., Zuñiga-Collazos, A., & Martínez-López, F. J. (2019). University Spin-Off: A Literature Review for Their Application in Colombia. *Journal of Environmental Management and Tourism*, *10*(1), 73–86. doi:10.14505//jemt.v10.7(33).08

Castrillón-Muñoz, A. J., Infante-Moro, A., Zuñiga-Collazos, A., & Martínez-López, F. J. (2020a). Capacities of the research groups at unicauca (Colombia) to develop spin-off-type undertakings. *Journal of Technology Management & Innovation*, *15*(1), 64–75. doi:10.4067/S0718-27242020000100064

Castrillón-Muñoz, A. J., Infante-Moro, A., Zúñiga-Collazos, A., & Martínez-López, F. J. (2020b). Generation of technological-based spin offs based on I+D+i results from the research groups of the University of Cauca Colombia [Generación de empresas derivadas de base tecnológica (Spin Offs), a partir de los resultados de I+ D+ i de los grupos de investigación de la Universidad del Cauca, Colombia]. *Información Tecnológica*, *31*(1), 67–78.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cerdán, Á. L. M. (2006). Análisis de la adopción de tecnologías colaborativas en Pymes. *Revista de economía y empresa*, *24*(56), 51-66.

Codara, L. (1998). *Le Mappe Cognitive*. Carocci Editore.

Curia, L., & Lavalle, A. (2011). Decision strategies in dynamic systems - applying fuzzy cognitive maps application to a socio-economic example [Estrategias de decisión en sistemas dinámicos - aplicando mapas cognitivos difusos aplicación a un ejemplo socio-económico]. *Journal of Information Systems and Technology Management*, *8*, 663–680. doi:10.4301/S1807-17752011000300008

Domingo-Carrillo, M. A., González-Rodríguez, R., & Chávez-Miranda, E. (2020). Identifying hotel revenue management implementation drivers. *International Journal of Information Systems and Software Engineering for Big Companies*, *7*(2), 33–48.

Ekong, U. O., Ifinedo, P., Ayo, C. K., & Ifinedo, A. (2012). E-commerce adoption in Nigerian businesses: An analysis using the technology-organization-environmental framework. In Leveraging Developing Economies with the Use of Information Technology: Trends and Tools (pp. 156-178). IGI Global.

Entorno turístico. (2020). *Artificial intelligence and hospitality, a good combination?* [La inteligencia artificial y el hotelería, ¿buena combinación?]. https://www.entornoturistico.com/inteligencia-artificial-hoteleria-buena-combinacion/

Europa Press. (2018). *Hoteliers are betting on the Internet of Things to personalize the experience* [Los hoteleros apuestan por el Internet de las Cosas para personalizar la experiencia]. https://www.europapress.es/turismo/hoteles/noticia-hoteleros-apuestan-internet-cosas-personalizar-experiencia-20180322111219.html

Feuerlicht, G., & Govardhan, S. (2010). Impact of cloud computing: beyond a technology trend. *Systems Integration*.

Flores-Fernandez, J., & Martínez-López, F. J. (2020). Historical cycles and prospective: Our future according to our past [Ciclos históricos y prospectiva: nuestro futuro según nuestro pasado]. *Revista de Pensamiento Estratégico y Seguridad CISDE*, *5*(1), 103–121.

García-Peñalvo, F. J. (2021). Digital Transformation in the Universities: Implications of the COVID-19 Pandemic [Transformación digital en las universidades: Implicaciones de la pandemia de la COVID-19]. *Education in the Knowledge Society*, *22*, e25465. doi:10.14201/eks.25465

García-Ruiz, R., & Pérez Escoda, A. (2021). Teacher's digital competence as key for strengthening a responsible use of the Internet [La competencia digital docente como clave para fortalecer el uso responsable de Internet]. *Campus Virtuales*, *10*(1), 59–71.

Gonçalves, M. J. A., Camarinha, A. P., Abreu, A. J., Teixeira, S. F., & Ferreira da Silva, A. (2020). An analysis of the most used websites in Portugal regarding accessibility web in the tourism sector. *International Journal of Information Systems and Tourism*, *5*(1), 19–28.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

González-González, C. S., Violant Holz, V., Infante Moro, A., Cáceres García, L., & Guzmán Franco, M. (2021). Educational robotics in inclusive contexts: The case of the hospital classrooms [Robótica educativa en contextos inclusivos: el caso de las aulas hospitalarias]. *Educación XX1*, *24*(1), 375-403. [] doi:10.5944/educxx1.27047

González-Zamar, M. D., Abad-Segura, E., & Belmonte-Ureña, L. J. (2020). Meaningful learning in the development of digital skills. Trend analysis. *IJERI: International Journal of Educational Research and Innovation*, (14), 91–110.

González-Zamar, M. D., Abad-Segura, E., & Gallardo-Pérez, J. (2021). Ubiquitous learning in arts education and visual languages: Trend Analysis [Aprendizaje ubicuo en educación artística y lenguajes visuales: Análisis de tendencias]. *Campus Virtuales*, *10*(1), 125–139.

Gutierrez, A., Boukrami, E., & Lumsden, R. (2015). Technological, organizational and environmental factors influencing managers' decision to adopt cloud computing in the UK.[. *Journal of Enterprise Information Management*, *28*(6), 788–807. doi:10.1108/JEIM-01-2015-0001

Iacovou, C. L., Benbasat, I., & Dexter, A. S. (1995). Electronic data interchange and small organizations: Adoption and impact of technology. *Management Information Systems Quarterly*, *9*(4), 465–485. doi:10.2307/249629

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020a). Key factors in the implementation of Cloud Computing as a service and communication tool in universities. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 631-636). 10.1145/3434780.3436698

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020b). Motivational factors in the insertion of digital skills in teaching. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 365-370). 10.1145/3434780.3436631

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020c). Motivational factors that justify the implementation of the internet of things as a security system in the hotel sector [Factores motivacionales que justifican la implementación del Internet de las Cosas como sistema de seguridad en el sector hotelero]. *Revista de Pensamiento Estratégico y Seguridad CISDE*, *5*(2), 81–91.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020d). The employment possibilities of the internet of things in the hotel sector and its training needs [Las posibilidades de empleo del Internet de las Cosas en el sector hotelero y sus necesidades formativas]. *Education In The Knowledge Society*, *21*, 14. doi:10.14201/eks.222643

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021a). Factors that influence the adoption of the Internet of Things in the hotel sector [Factores que influyen en la adopción del Internet de las Cosas en el sector hotelero]. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, (E41), 370-383.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021b). The acquisition of ICT skills at the university level: The case of the Faculty of Business Studies and Tourism of the University of Huelva [La adquisición de competencias TIC en el ámbito universitario: El caso de la Facultad de Ciencias Empresariales y Turismo de la Universidad de Huelva]. *Revista de Medios y Educacion*, (60), 29–58.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021c). Key Factors in the Implementation of the Internet of Things in the Hotel Sector. *Applied Sciences (Basel, Switzerland)*, *11*(7), 2924. doi:10.3390/app11072924

Infante Moro, J. C. (2017). *User perception to improve the use of Social Networks as a communication channel in the hotel sector* [Percepción de los usuarios para la mejora del uso de las Redes Sociales como canal de comunicación en el sector hotelero] (Doctoral Thesis). University of Huelva.

Kula, V., & Tatoglu, E. (2003). An exploratory study of Internet adoption by SMEs in an emerging market economy.[. *European Business Review*, *15*(5), 324–333. doi:10.1108/09555340310493045

Laforet, S. (2011). A framework of organisational innovation and outcomes in SMEs. *International Journal of Entrepreneurial Behaviour & Research*, *17*(4), 380–408. doi:10.1108/13552551111139638

Lin, A., & Chen, N. C. (2012). Cloud computing as an innovation: Percepetion, attitude, and adoption. *International Journal of Information Management*, *32*(6), 533–540. doi:10.1016/j.ijinfomgt.2012.04.001

Lippert, S. K., & Govindarajulu, C. (2006). Technological, organizational, and environmental antecedents to web services adoption. *Communications of the IIMA*, *6*(1), 14.

Maridueña, M. R., Leyva, M., & Febles, A. (2016). Modelado y análisis de indicadores de ciencia y tecnología mediante mapas cognitivos difusos. *Ciencias de la Información*, *47*(1), 17–24.

Martínez-López, F. J., García-Ordaz, M., Arteaga-Sánchez, R., & Infante-Moro, A. (2015). The presence of large Spanish companies in online social networks. *Journal of Marketing Analytics*, *3*(4), 171–186. doi:10.1057/jma.2015.15

Martínez Pérez, S., Fernández Robles, B., & Barroso Osuna, J. (2021). Augmented reality as a resource for training in higher education [La realidad aumentada como recurso para la formación en la educación superior]. *Campus Virtuales*, *10*(1), 9–19.

Mehrtens, J., Cragg, P. B., & Mills, A. M. (2001). A model of Internet adoption by SMEs.[. *Information & Management*, *39*(3), 165–176. doi:10.1016/S0378-7206(01)00086-6

Morales Salas, R. E., Infante-Moro, J. C., & Gallardo-Pérez, J. (2019). Mediation and interaction in a VLE for effective management in virtual learning [La mediación e interacción en un AVA para la gestión eficaz en el aprendizaje virtual]. *Campus Virtuales*, *8*(1), 49–61.

Morales Salas, R. E., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020). Evaluation of virtual learning environments: A management to improve. *IJERI: International Journal of Educational Research and Innovation*, (13), 126–142.

Morales Salas, R. E., & Rodríguez Pavón, P. R. (2020). Digital ICT skills applied in organizations [Las competencias digitales en TIC aplicadas en las organizaciones]. *International Journal of Information Systems and Software Engineering for Big Companies*, *7*(1), 25–35.

Mouratiadou, I., & Moran, D. (2007). Mapping public participation in the Water Framework Directive: A case study of the Pinios River Basin, Greece. *Ecological Economics*, *62*(1), 66–76. doi:10.1016/j.ecolecon.2007.01.009

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Özesmi, U., & Özesmi, S. L. (2003). A participatory approach to ecosystem conservation: fuzzy cognitive maps and stakeholder group analysis in Uluabat Lake, Turkey. *Environmental Management*, *31*(4), 518-531. doi:10.100700267-002-2841-1

Özesmi, U., & Özesmi, S. L. (2004). Ecological models based on people's knowledge: A multistep Fuzzy Cognitive Mapping approach. *Ecological Modelling*, *176*(1), 43–64. doi:10.1016/j.ecolmodel.2003.10.027

Pacheco-Cortés, A. M., & Infante-Moro, A. (2020). ICT resignification in a virtual learning environment [La resignificación de las TIC en un ambiente virtual de aprendizaje]. *Campus Virtuales*, *9*(1), 85–99.

Pangrazio, L., & Sefton-Green, J. (2021). Digital Rights, Digital Citizenship and Digital Literacy: What's the Difference? *Journal of New Approaches in Educational Research*, *10*(1), 15–27. doi:10.7821/naer.2021.1.616

Papageorgiou, E. I., Markinos, A. T., & Gemtos, T. (2009). Application of fuzzy cognitive maps for cotton yield management in precision farming. *Expert Systems with Applications*, *36*(10), 12399–12413. doi:10.1016/j.eswa.2009.04.046

Papageorgiou, E. I., & Salmerón, J. L. (2013). A Review of Fuzzy Cognitive Maps Research During the Last Decade. *IEEE Transactions on Fuzzy Systems*, *21*(1), 66–79. doi:10.1109/TFUZZ.2012.2201727

Premkumar, G., Ramamurthy, K., & Crum, M. (1997). Determinants of EDI adoption in the transportation industry. *European Journal of Information Systems*, *6*(2), 107–121. doi:10.1057/palgrave.ejis.3000260

Premkumar, G., & Roberts, M. (1999). Adoption of new information technologies in rural small businesses. *Omega*, *27*(4), 467–484. doi:10.1016/S0305-0483(98)00071-1

Recio Muñoz, F., Silva Quiroz, J., & Abricot Marchant, N. (2020). Análisis de la Competencia Digital en la Formación Inicial de estudiantes universitarios: Un estudio de meta-análisis en la Web of Science. *Píxel-Bit, Revista De Medios Y Educación*, *59*, 125-146. [] doi:10.12795/pixelbit.77759

Rogers, E. (2003). Diffusion of innovations. Delran, NJ: Simon & Schuster.

Solana Gutierrez, J., Rincón Sanz, G., Alonso González, C., & Garcia De Jalon Lastra, D. (2015). Use of Maps of Diffuse Knowledge (FCMs) in the prioritization of river restoration: Application to the Esla River [Utilización de Mapas de Conocimiento Difuso (MCD) en la asignación de prioridades de la restauración fluvial: Aplicación al río Esla]. *Cuadernos de la Sociedad Española de Ciencias Forestales*, *41*, 367–380.

Sonnenwald, D. H., Maglaughlin, K. L., & Whitton, M. C. (2001). Using innovation diffusion theory to guide collaboration technology evaluation: work in progress. In *Proceedings Tenth IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises. WET ICE 2001* (pp. 114-119). IEEE. 10.1109/ENABL.2001.953399

Sosa Sierra, M. D. C. (2011). Artificial intelligence in business financial management [Inteligencia artificial en la gestión financiera empresarial]. *Revista científica Pensamiento y gestión*, (23), 153-186.

Tecnohotel. (2018). *Artificial Intelligence, advantages and challenges of automating the hotel industry* [Inteligencia Artificial, ventajas y retos de automatizar la industria hotelera]. https://n9.cl/lyia

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

To, M. L., & Ngai, E. W. (2006). Predicting the organisational adoption of B2C e-commerce: An empirical study. *Industrial Management & Data Systems*, *106*(8), 1133–1147. doi:10.1108/02635570610710791

Tu, M. (2018). An exploratory study of Internet of Things (IoT) adoption intention in logistics and supply chain management. *International Journal of Logistics Management*, *29*(1), 131–151. doi:10.1108/IJLM-11-2016-0274

Viana, J., & Peralta, H. (2021). Online Learning: From the Curriculum for All to the Curriculum for Each Individual. *Journal of New Approaches in Educational Research*, *10*(1), 122–136. doi:10.7821/naer.2021.1.579

Weber, R. H. (2010). Internet of Things–New security and privacy challenges. *Computer Law & Security Review*, *26*(1), 23–30. doi:10.1016/j.clsr.2009.11.008

KEY TERMS AND DEFINITIONS

Adjacency Matrix: In this study, a square matrix that is used as a shape to represent cause-effect relationships based on the opinion of each of the experts. These relationships are quantified by values of the interval [-1,1].

Causal Map: Graph obtained from the cause-effect analysis of the experts on the different factors involved in decision-making.

Collective Cognitive Matrix: In this study, a square matrix that is used as a form to represent cause-effect relationships based on the opinion of the experts interviewed. The assessment of each of the experts is collected and, subsequently, the average assessment of each of these relationships is recorded (quantified by averages).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 15
Peruvian Consumer Responses to the Use of Technologies in the Context of COVID-19

Yezelia Danira Caceres Cabana
https://orcid.org/0000-0002-5027-8081
Universidad Nacional de San Agustín de Arequipa, Peru

David Aguilar del Carpio
https://orcid.org/0000-0001-9911-6537
Universidad Nacional de San Agustín de Arequipa, Peru

Erika Velásquez Chacón
https://orcid.org/0000-0003-2247-3100
Universidad Católica San Pablo, Peru

Juan Mardonio Rivera Medina
https://orcid.org/0000-0003-0627-7871
Universidad Continental Perú, Peru

ABSTRACT

The pandemic caused by COVID-19 has led many states, in an attempt to control the spread of the virus, to decree social immobilization, which meant the introduction of restrictions on the free movement of people and the opening of shops. This led them to seek new marketing channels for purchases. Among these, ICTs have been important. This is the focus of the analysis in this document. Through surveys and interviews, information was obtained, divided into four age groups, which showed that an important part of the population has had to resort to ICTs to acquire goods and/or pay for services. This change in the way of acquiring had different particularities according to the age group analyzed, with a greater change in the oldest group (56 to 74 years old). It can be concluded that the massification of these tools has generated a change in the ways of acquiring products, and this is likely to transcend the pandemic and these channels will be maintained and strengthened in the future.

DOI: 10.4018/978-1-7998-6985-6.ch015

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

In Latin America, the assimilation of technology in productive processes and supply chains is still limited, and there is also a digital divide, presenting this as an obstacle for important sectors of the population, affecting rural areas to a greater extent because they have greater limitations in access to and use of digital technologies; situations that have had a strong impact on the relationships between individuals, creating new means through which they interact with their environment in different ways. Some of these spaces are the ways in which people consume and/or acquire the goods necessary to satisfy their physical and/or spiritual needs. Therefore, "digital infrastructure is a fundamental component in maintaining economic resilience" (CAF Development Bank of Latin America, 2020, p.17).

The pandemic caused by COVID-19 has shown that the Peruvian state and its population were not prepared to face crisis situations at this level, showing a strong vulnerability in many sectors. The importance of this chapter lies in the fact that it provides an understanding of the dynamics that exist in developing economies, where a large number of the population still does not have access to or feels culturally distant from the use of digital tools for consumption and/or the sale of products. All this information will serve the sectors involved in the planning and elaboration of public policies that will help strengthen the implementation of activities and provide better conditions for value generation.

This analysis is based on the city of Arequipa, which "is the second pole of industrial development in the country, highly diversified with a base made up of leading companies producing consumer goods, inputs and capital goods of national and regional scope" (Banco Central de Reserva del Perú, 2016, p.13). Regarding the economic aspect, in this same department 95.8% of the companies are "Micro" (INEI, 2014), most of which have been affected by the isolation measures imposed by the government to face COVID-19.

In this context, companies have not been able to carry out their productive activities normally, which has led to a significant drop in their income. This socio-economic situation will have a direct impact on the increase of precarious employment, informality and poverty levels. As a result, many people who were entering and establishing themselves in the emerging middle class could return to poverty. This would add more people to the informal sector of the economy, which already accounted for more than 70 per cent before the pandemic. The largest informal sector is the commercial sector, which is seen as a refuge for workers in other sectors. All this shows that these people and businesses will have to reinvent themselves commercially to survive, just as consumers have to connect to new channels to acquire the goods they need.

It is possible that the effects of COVID 19 have drastically changed the purchasing habits and mechanisms of Arequipa's consumers, which is why this document aims to micro-analyse consumers' responses to the use of information and communication technologies. In this sense, we assume that this has required, on the part of consumers and producers, the use of a set of information on the proximity, quality and availability of products, and of the universe of possible consumers on the producers' side as well. We also consider the tendency to replace unsophisticated information equipment with devices that allow access to networks and web pages, together with access to an internet connection that facilitates the connectivity of the agents involved.

The research carried out is non-experimental, exploratory and descriptive, with the use of qualitative and quantitative tools, through the application of convenience sampling, aimed at people living in the city of Arequipa, Peru, of different age groups classified as follows: between 10 - 24 years old, 25 - 40 years old, 41 - 55 years old, 56 - 74 years old.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the combined information between qualitative and quantitative methods, we tried to obtain information on the access and use of the internet through social networks to search for or purchase a product before and during COVID 19, for this purpose, 237 surveys were carried out at a quantitative level using an electronic form and 25 in-depth telephone interviews which were recorded and transcribed in their original language in Spanish and then translated into English by the authors, this information allowed us to know and understand the thinking and logic of the people.

For the selection of participants, sampling started with families known to the researchers and was expanded using the snowballing technique. The surveys and interviews were conducted between November and December 2020 by telephone.

BACKGROUND

Since the beginning of the pandemic, changes in consumer behaviour have been generated, starting with the stockpiling of goods, however, social isolation measures limited the availability and accessibility of products and services (Casco, 2020, pp. 98-105). At the same time, the population sought to make the best use of their resources, due to the feeling that the situation could become increasingly complex, and the fear of the latter led to an increase in their inventories to avoid shortages.

From the above, it can be stated that consumers learn and have new habits, now the shop comes home and this is because "All consumption is location- and time-dependent. Consumers develop habits over time about what to consume, when and where" (Sheth, 2020, pp- 280-283). Kirk and Rifkin (2020) have examined behaviours in three phases of first reacting, then coping and adapting in unusual patterns of behaviour in the face of the COVID-19 pandemic. It is therefore evident that the population modifies its behaviour and this behaviour is likely to persist over time.

The Development Bank of Latin America [CAF] (2020) in the report: El estado de la digitalización de América Latina frente a la pandemia del COVID-19, presents the importance of digitalisation as mitigating factors of the disruption of the pandemic, where digital technologies have allowed counteracting isolation and facilitating the functioning of economic systems, where the importance of digitalisation is shown. This is similar to consumers in Arequipa, where the use of information and communication technologies has supported families, mainly in urban areas and with limitations in rural areas. In this context, the use of information and communication technologies has supported and facilitated many of the needs of the population in education, health, teleworking, and connecting with family members, among other simultaneous uses, although with limited access, due to increasingly precarious economic levels, job losses or a decrease in production and sales at the business level.

According to CAF (2020), Latin America has positioned itself at an intermediate level in its digital ecosystem, mitigating the effects of the pandemic, although Latin American networks have also been affected by the exponential increase in traffic, although with effects on the decrease in broadband speed.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

FRAME OF REFERENCE

Consumption

The processes involved in the actions of buying, using, searching for, evaluating and discarding goods and services in order to satisfy consumer needs have been addressed not only by economics, but also by psychology, sociology and anthropology. This has gone from the evaluations revealed in a first approach on the rationality of the consumer to then broaden in its interpretation towards the elements that give impulse to some behaviours according to the environment, progressively giving way to the psychological emphasis and neuroscience.

From an economic perspective, Varian (2014) analyses consumer preferences at a more general level, indicating that this requires "not only a complete list of goods that could be consumed, but also a description of when, where and under what circumstances they could be obtained".

On the same subject, but from an anthropological perspective that seeks to approach the cultural behaviour of human beings in terms of their consumption decisions, Caceres Cabana et al. (2020), state that consumption must be understood from its simplest and most dispossessed dimension, to the level at which real political, economic and even cultural interests are effectively hidden behind simple buying and selling actions. Although individuals have to consume in order to subsist, consumption is no longer only subject to the idea of subsistence, but is linked to other factors, such as prestige, taste, etc.

In this sense, Akerlof and Shiller (2015) introduce the idea that consumers actually buy things they do not need, in a state of manipulation of their senses by advertising and the social environment, generating a need where a false scarcity is created, which will condition or encourage unnecessary consumption. Regarding how scarcity focuses people's senses, Mullainathan and Shafir (2016) indicate that "Scarcity captures the mind. Just as hungry subjects thought about food, when we suffer from scarcity of any kind, it absorbs us. The mind automatically and powerfully focuses on unsatisfied needs". Likewise, following the previous authors, they argue that scarcity is not only a state of mind due to the frustration of having less or little, but rather, it will influence how we perceive things, imposing itself in the mind over other activities or needs. This is why many population groups respond in a culturally differentiated way to structure their consumption needs. Thus, consumption can be seen not only from a purely needs-based aspect, but also in socio-cultural terms.

Evidence from Hao et al. (2020) reveals that fresh food e-commerce channels are more likely to be associated with panic due to a higher likelihood of supply shortages.

Consumers

Digital tools were presented as those that could somehow link the consumer with the supplier, opening the possibility of continuing to maintain marketing channels in the face of the difficulties described above, this phenomenon with the pandemic has accelerated the use of many of these digital media, due to the need of many families who, despite being in a situation of compulsory social isolation, needed to continue acquiring goods and/or paying for services. Therefore, technological resources have become a necessity for consumers in a context of limited financial and digital inclusion, insufficient use and massification of means of payment and insufficient access to information technologies.

Compulsive purchases showed changes in consumption behaviour in the face of the pandemic; at the beginning there was hoarding of products, generating shortages, but this was later controlled. In

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

this regard, the literature shows that there is a relationship between conscious thought and unconscious processing as elements that lead to human behaviour (Masicampo & Baumeister, 2011).

The generational review to understand consumer behaviour is complex because each one has particularities and different characteristics in the context of each context, where the generations of 30 years old are considered as the time of progress, because they have the ability to create and those of 60 years old in the stage that diminishes their role in public life, considering "aspects such as communication, use of technology, motivation, recruitment, incentives, among others", for Diaz-Sarmiento et al. (2017) generates challenges for their understanding, because they adopt particular positions in the way of thinking and acting.

The theoretical foundation reported by Díaz-Sarmiento et al. (2017), determines:

a) Baby Boomers, the 50-70 year olds born between 1946 and 1964, are characterized by their dedication and even workaholism. Empowered and looking forward to the best in life, they are a generation concerned with the pursuit of status, loyalty and quality of life.
b) Generation X, which refers to those born between 1965 and 1977, who hold middle and senior management positions and are the parents of Millennials and Centennials and were at the center of the consumerism of the 1980s. This population grew up where both parents worked or were divorced, influenced by events such as the emergence of personal computers and the expansion of the internet.
c) Generation Y, Millennials, are those born between 1980 and the beginning of the 21st century, are children of the last Boomers and the first X and grew up in a culture of protected and loved children, characterized by the use of technology as part of their lifestyle. They have grown up with the internet, smartphones, accelerated technological advances, social networks and with these, instant information. On the other hand, Millennials have had so many awards and trophies growing up that many of them think they should be promoted in their jobs every two years regardless of their performance.
d) Generation Z, born since the beginning of the 21st century, stand out for their assumed capacity for rapid response, their desire for immediacy and continuous interaction, and are also conceived as experts and competent in ICT. In this sense, the preferences of this group lean towards visual information and easy performance in digital and visual environments where several tasks are managed at the same time, a phenomenon is known as multitasking (Bennett et al., 2008; Gallardo, 2012; and Fernández & Fernández, 2016; Cassany & Atalaya, 2008, Reig & Vilchez, 2013 cited in Pérez-Escoda, Castro-Subizarreta, & Fandos-Igado, 2016).

Culture and Technology

The cultural issue, in times of globalization, sometimes goes unnoticed. However, different societies still behave differently in response to certain stimuli or situations. The cultural factor could explain this, which we understand as "the traditions and customs, transmitted through learning, that shape and guide the beliefs and behaviour of the people exposed to them" (Marcus & Fisher, 1999). (Phillip Kottak, 2011). In the same vein, it is important to make reference to the fact that "the cultures of the world's people need to be constantly rediscovered and they reinvent themselves in changing historical circumstances". (Marcus & Fisher, 1999)In other words, they are subject to environmental pressures which result in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

processes of adaptation, not only biological, like all animals, but also in processes of cultural adaptation, which is inherent to human beings.

Regarding technological advances, Ogbum & Thomas (1922) state that society, as well as its norms and values, will change more slowly than technology, so the cultural change will also change. In this sense, on the cultural link with technologies, Thomas et al. (2005) indicate that there are several social factors that will influence the way in which this takes place, pointing out that the degree of homogeneity and level of equality of societies, the degree of communication with the outside world (especially influenced by migration), the experience of gender and ethnicity in different societies, cultures of different generations, religion, education and literacy and language, will condition the way and speed in which societies adapt to technology.

Information and Communication Technologies (ICTs)

For this study, information and communication technologies (ICTs) are understood as the information technology tools used by consumers in email services, web use, data storage, electronic banking, virtual communities managed through devices such as computers, mobile phones and tablets. The World Economic Forum (2019) suggests that ICTs are considered to be the set of tools related to the digitised transmission, processing and storage of information.

It is also a facilitating element of entrepreneurship in a globalised context, the contributions of Serrano et al. (2018) highlight that through the use of electronic devices and the inclusion of computer programs, it is possible to establish channels of participation between different communities, groups and collectivities, streamlining processes, times and costs; ICTs are considered as a mediating tool between the processes of business activities, enabling external relations between customers, suppliers, markets, potential consumers, among others; and through their operation they create improvements in the service or product offered, increasing the efficiency of trade and its economic growth.

In relation to the opportunities that organizations gain by incorporating ICTs in their activities, according to Martelo et al. (2017, pp. 87-94), advantages can be expressed in the streamlining of processes, cost savings, reduction of personnel, likewise Martínez Coral (2017, pp. 63-72) highlights its effectiveness in the promotion of education, a very important aspect for individual and collective management, additionally Serrano et al. (2018) attributes it a potential to generate value, disseminate knowledge, improve quality of life. These elements are relevant in a context of accelerated changes in daily and social life caused by the pandemic where it can be seen that the behavior and use of this technology have become widespread and has generated business opportunities and greater profitability in business organizations linked to mobile telephony, computers, internet, digital television, applications, etc., a situation that has involved business organizations.

The adoption and use of ICTs by market players depending on the internal capacities, size, and culture of the organizations; an aspect that is linked to e-commerce, when goods and services are bought and sold through the internet, making it possible to process information from commercial transactions to create, transform and redefine relationships between organizations and individuals, as the case may be, in order to create value. There is an evolution in e-commerce through the use of the World Wide Web in part or all of the transaction cycle supported by a range of technologies ranging from mobile devices (m-commerce), social networks (social commerce), email and others (Jones et al., 2013, pp. 164-175).

On the effects of using decision-inducing user interface design elements to guide consumer behaviour in digital choice environments; evidence reveals that different stimuli and more versatile adaptation

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to users based on the data it provides could have notable implications for guiding user behaviour in a desirable way under a choice architecture towards related items that influence purchase by motivating additional consumption that was not originally planned by the user (Weinmann et al., 2016, pp. 433-436).

MAIN FOCUS OF THE CHAPTER

The use of ICTs in Arequipa

According to the results found through this research on the interaction experiences of the different age segments with ICTs to carry out search procedures and purchase goods and services before and during the pandemic, it can be seen that the most active segment that has been exploring and applying ICTs in commerce is mainly concentrated in the age segment corresponding to generations X, Y and Z, as well as in the preference of purchases.

Generation Z and Generation Y reported frequent use of social networks and websites, compared to 8% and 3% for Generation X and baby boomers respectively; with this experience extending to shopping and searching for commercial transactions via social networks for 37% of Generation Z and 16% of Generation Y.

On the other hand, to see the projection regarding the development of the perspective and sustained growth in the use of ICTs by the different age segments analyzed, 35% of generation Z and 18% of generation Y state that they will continue to use ICTs for commerce in the post COVID 19 pandemic scenario, compared to 6% of generation X and 2% of baby boomers respectively, This situation suggests that the growing trend of supply and demand decisions for goods and services through ICTs will be irreversible, and that connectivity conditions must be improved to ensure the efficiency and effectiveness of digital market players.

Regarding the use of ICT applied to commerce, it is also worth highlighting the predominance of the use of devices, highlighting that Generation Z and Generation Y are the ones that dominate in their use: 29% and 17% use mobile phones, 19% and 5% use laptops, 7% and 4% use fixed computers, in this case it is worth paying particular attention to the behaviour of Generation X and the Baby Boomers who stand out more in the use of mobile phones with 8% and 6% respectively, In this case, it is important to pay particular attention to the behaviour of Generation X and Baby Boomers, who use laptops (8% and 6% respectively), while 2% of both age segments do not use laptops, and the use of fixed computers is not relevant, which indicates that the age segments analysed generally prefer the technological sense of portability, simplicity, accessibility and price, which is expressed in the privileged use of mobiles.

The use of ICTs has brought about changes in consumer behaviour, with consumers accessing more information through the internet and social networks about the products and services they want to buy. These changes have been accelerated by the pandemic, and the data from Arequipa show that people in the Y and Z generation age groups are those who use ICTs frequently, reaching 27% and 13% respectively; Both segments explain the new consumer behaviour and the use of modern marketing channels, while there is a natural resistance among people in the 41-74 age groups, who use ICTs less than 6%. In this respect, an interviewee in the 25-40 age group says: "*Normally if you can't find something on one website you can look on another, so I always find what I'm looking for. I look first at how many people have shopped there and what comments they have left on this shop or business. Once I have done that I see if they are trustworthy. (interviewee 13)*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Consumer Behaviour by age Group During the Pandemic

Hereafter, to describe consumer behaviour towards technology, this will be understood as the use of mobile or computing devices for the acquisition or selection of goods and services. The analysis is carried out according to the age groups referred to in the reference framework, which describes consumers differentiated by generational groups.

Centennials or Generation Z

The people in this group are what we could consider as digital natives, i.e. those who were born in a recent era, which has allowed them to be more familiar with the aforementioned technologies, considering them as something every day and normalized, very different from those who had to go through a learning process for their use, the latter are known as digital migrants.

Based on the results found, it can be seen that the majority of people belonging to this age group have used digital tools to make their purchases during the pandemic, this being a behaviour that already existed before the pandemic in almost all respondents, so it did not represent any shock in the way of accessing the goods necessary for their consumption, as stated in the statement: *If I use social networks to buy, I think it is an option because it allows me to not need to leave the house. (interviewee 6)*

It can be added that the use of digital means of payment such as credit cards is not widespread among this age group, as their financial inclusion is still low. This is due to the fact that many are still in education and have not yet obtained a job or financial independence, so they are not creditworthy. This is reinforced by the evidence found in the 2012 Pilot Survey on Access to and Use of Financial Services, prepared by the Peruvian Superintendency of Banking and Insurance, when analysing the client profile within the financial system, which indicates that "Banking penetration (having an account, credit or credit card) had a slight bias towards women, who accounted for 52% of the total number of clients in the financial system. By age group, the results suggest that the financial system is concentrated among people between 26 and 50 years old (64%), followed by clients over 50 years old with 26%, while young clients between 18 and 25 years old represented a lower percentage (10%)" (Superintendencia de Banca, Seguros y AFP - BID, 2012).

Although most people in this age group had made purchases before, it is believed, as mentioned above, that the factor of using digital means of payment such as credit cards conditions virtual purchases, which is why 30% of those who said they had ever made online purchases indicated that they had only done so since the pandemic. It is therefore not surprising that almost 40% of this group said that they paid in cash when shopping online.

Generation Y

The people in this generation have a very similar familiarity with the use of technologies to the age group analysed above, since, although most of them were not born or raised with the existence of this type of digital tool, this group has had early contact with them, so it is not surprising that 65% of respondents belonging to this consumer group stated that they were already in the habit of using them to purchase consumer goods before the pandemic. This is further strengthened by the evidence that this age group has almost all shopped online, with nearly 85% indicating that they had done so at some point in the survey, which reinforces the hypothesis put forward earlier that they would have a strong link to the use of these

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

tools. Of this group, only 29% started making such transactions in the wake of the pandemic, with the increase occurring mainly between June and August 2020. These purchases were mostly associated with food consumption. Also, almost half of the respondents indicated that they had had problems acquiring basic necessities from April to the date of this study. It is important to note in this group that 65% do not make their purchases with support, thus reinforcing the statement above that they have an adequate use of technology. This is why they consider that over time the technology itself has become more user-friendly, as reflected in the response of 48% of the respondents. However, we found in the qualitative analysis that there are members of this group who still show distrust, as indicated in the interview: *I don't use it, I think it's a deception, they can put pretty pictures and deliver other things. (interviewee 1)*

Generation X

The penultimate group analysed is Generation X. Here we can already find a slightly more marked differentiation with the age groups described above, since, although many of the people in this age range are far from a total link with the digital world, most of them have had experiences or are aware of the existence of these tools as possibilities for accessing consumer goods. Cullen (2001) indicated that, in terms of "ICT use, there are specific groups of people who are at a disadvantage, such as people with low incomes, low levels of schooling, the unemployed, older adults, residents of rural areas, people with disabilities, women and girls". Thus, in some cases, the consumption behaviour of this group during the pandemic was modified by the need to be able to acquire the goods necessary for their subsistence. However, in some cases, they showed distrust, seeing it as a traumatic event, because it represented new experiences and/or entering an unknown and disconcerting scenario by leaving aside the direct transaction between buyer-seller, which is reinforced in the idea of trust and visualization of the good in real-time and not in the idea of paying for or requesting a product for which one does not necessarily have all the information about its quality and characteristics requested. In relation to what has been argued in the previous lines, Martínez Domínguez (2018) indicates that older people are less adapted to new technologies than younger people.

A relevant finding for people in this group is that of the total number of people who indicated that they used social media to make purchases, 42% indicated that they had not made purchases through social media or other digital channels prior to the pandemic. This shows that a significant segment of the population changed their behaviour or had to adjust their behaviour due to the restrictions mentioned above. Also, all of the above-mentioned people indicated that they found social networks more user-friendly, which may indicate that there was a cultural and/or subjective factor that disengaged people in this group from earlier familiarisation with the use of digital media. *In* this sense, a study on technologies in the lives of older adults indicates that for people in this age range, "Some uses are incipient, such as economic transactions, due to a certain distrust and lack of transparency [that] older people perceive", reinforcing their idea by indicating that in older people, the traditional approach to ageing creates real barriers that prevent them from visualising the potential or facilities that the use of digital technologies can bring for the improvement of the quality of life of older people (Rivoir et al., 2019, p. 310).

The Baby Boomer Generation

The group of citizens is the most detached from the use of digital tools, due to the fact that they were born and developed their first skills in a social scenario in which these did not exist, so they are accustomed

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to the use of analog objects, placing them culturally more distant from the naturalization of the use of these devices. In this sense, Martínez Domínguez (2018) indicates that *people in the 19-30 and 31-50 age groups are more likely to use these technologies, compared to people over 50.* This has undoubtedly had an impact on how they conceive their use, and can change, as mentioned in previous paragraphs, the feeling of distrust and impossibility of approaching consumption channels to a completely different one that shortens gaps and generates technological learning options. It is therefore not surprising that the use of these technologies has been distant and complicated in the pandemic. According to the results found, this population group had little or no use of digital channels for the acquisition of consumer goods before the pandemic. However, as a result of COVID-19, this has had to change according to the needs that arose, as can be seen in the following statement: *I was also looking at clothes on the internet and I don't know if they will arrive and I don't know, there is always that doubt that they will not arrive because we are not used to ordering on the internet. (interviewee 2)*

The findings found in this age group show that of the total number of people surveyed, 67% have not used social networks to purchase any product. Of those who have used social networks, 47% have only done so since the pandemic, suggesting that these people have modified their behaviour due to the needs generated by social isolation. However, they still show difficulties in the use of digital means of payment, as they indicated that 75% of their payments are made in cash, i.e. the transaction was made on delivery. In addition, all of the people in this group used a family member's social network to make the purchase. The data collected show that one of the factors why people in this age group do not make purchases through these means is a cultural issue, which is demonstrated by the distrust they have indicated towards this type of non-face-to-face transaction. It is important to highlight that the highest percentage of their purchases were concentrated in food.

Thus, in this group, there is evidence of serious difficulties in the use of social networks, at least not on a constant basis, the lack of digital means of payment by these people and the disengagement with the management of digital tools, which may be known but mostly used in a basic way without exploiting all its functions. Thus, in most cases, people of these ages had to ask for support or be assisted by a younger person, who through advice or the use of their own networks facilitated and/or enabled the acquisition of goods or services that were usually used in the pre-pandemic scenario. This suggests that a certain level of dependency of people in this group on other age groups may have been generated. It is important to point out that, according to the information gathered, the confinement caused by the COVID-19 pandemic has also allowed people in this age range, due to the need to make purchases, to generate levels of learning that are important in terms of reinforcing confidence in these media, as can be seen in the statement: "*I have only started to use Internet shopping with the pandemic" (interviewee 3).*

Of the total number of interviews conducted, more than 70% indicated that they have ever used or purchased through social media. The rest of the respondents said that they were wary of making purchases through these channels. In this regard, it is worth noting that, of the total number of interviewees, one-third have only started using social networks since the pandemic, while 15% have increased their use, either as a means of social networking or as an instrument for making purchases.

Finally, the majority of people, around 85%, responded that, through websites or virtual media, they find what they are looking for, which represents an interesting result in terms of the variability of product offerings that guarantee a connection with the end consumer, generating bonds of trust in terms of being able to satisfy their needs. In terms of payment methods, almost all respondents indicated that they had or had used digital payment methods, such as credit cards, use of bank applications, deposits with agents, etc.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

This reinforces the idea that there has been a change in behaviour as a consequence of the pandemic, meaning that people, faced with the need to be able to purchase the goods they used to use in the pre-pandemic scenario, have had to use digital tools, not only to search for products, but also to pay for them.

Among the factors that keep many sectors of the population away from the use of these tools are cultural, educational, distrust in the use of apps and lack of knowledge of their management, for example shopping online was easy for young people but not for adults and older adults who are less close to technology so it is necessary to empower citizens to make proper use of these resources, adults have the need to have contact with networks and learn from them to get closer and access them, without any dependence, in a friendly, agile and safe way. In this sense, virtual environments should facilitate processes with virtual instructions to access and make use of the services that are provided, due to the fact that the teaching levels of third parties are difficult, as it is observed that, if young people are closer to technology, it is not necessarily possible to find the availability to teach or help their relatives or close people because they have different interests and do not have the time.

Marketing Channels

Information and communication technologies ICT as an innovation perspective have considerably modified geography and barriers as influential elements in the dynamics of all types of marketing channels, this situation has influenced changes in the commercial and financial management of economic agents, at the same time the flow of abundant information to the consumer gives him greater power in his ability to choose; including the retail sector, thus constituting a process of commercial innovation based on internet support changing in shop approach, distribution channels and payment methods mainly (Garcia & Gil-Saura, 2016).

Research contributions in other geographical contexts reveal that in the COVID-19 environment, wholesale and emerging channels grew, achieving 63% and 57% share for each (Kantar 2020), adding this change to e-commerce and telesales that already showed a positive trend, in Peru there was an increase in purchase tickets in e-commerce platforms that went from 45 to 84 soles and telephone orders that increased from 10 to 21 soles (Peru Retail, 2020). With the prominence of receiving orders via telephone and WhatsApp, as well as the incorporation of less physical contact payments, accompanied by sanitary measures.

In Arequipa, the arrival of Covid-19 has configured a new scenario that was not foreseen, both in modern and traditional channels (BBVA Research, 2016), the latter represented mainly by markets, bodegas, small pharmacies, hardware stores, market stalls, kiosks with more points of sale, closer, more accessible and personalized.

On the other hand, in a crisis, decisions that in normal times could take years of deliberation are approved in a matter of hours (Harari, 2020). This is evident in the creativity of entrepreneurs who, having lost their jobs, initiate a variety of informal economic activities using social networks as an ally. An example of this is the emergence of informal mobile food trade by entrepreneurs motivated by scarcity and the need to take advantage of sales opportunities (Malone et al., 2020). This is a natural process brought about by emergencies, where progress is rapid.

One opportunity that is emerging in the midst of the pandemic is the delivery service, which is currently not fully developed, but under the adversities presented and the fear of contagion, business initiatives in this area are emerging. In the COVID-19 context, commerce has expanded by making use of ICTs as it has no geographical or time limits, where its implementation is more focused on consolidat-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing the loyalty of the new digital consumer, who in turn influences other potential consumers through their opinions and recommendations, which become a decisive factor in the success or failure of a given business model. For this, commerce has to manage digital marketing for the interconnection with the consumer through a digital sales channel, with a website positioned in social networks, an e-commerce platform established as a virtual shop; vigilant of the expectations of the digital customer that allows to knowing the new consumer habits and trends of the network that causes a different and authentic experience around a product or service, valuing personalization in the purchase by comparing prices instantly satisfying demands in real-time where email plays the role in customer loyalty (Junta de Andalucía, 2020).

The pandemic has generated different reactions from the business point of view, one of them is the reactivation and reinforcement of e-commerce platforms to respond to the confinement. However, the implementation of service models such as D2C (Direct to consumer), which was already developed but is going through a process of consolidation necessary to face the current situation, is beginning.

On the other hand, we can see that, in the face of all this, there is a vulnerability in the different markets, so marketing channels have had a significant impact on the pandemic, in terms of the adaptation process in the future of sales. However, the consequences for customers and the impact on the economy remain uncertain.

Issues, Controversies, Problems

In the context of the pandemic, consumption needs to be addressed by trying to better understand people's decision-making process under uncertainty, giving rise to the analysis of behavioral economics to explain that people's consumption decisions are also motivated by cognitive, social and emotional aspects of humans, filling the gap of the rational agent approach.

As mentioned above, the pandemic caused by COVID-19 has forced many people, either voluntarily or by government order, to stay at home and thus to disengage from the consumption channels to which they were accustomed, modifying not only habits, but also the forms and goods to which they had the possibility of accessing. Now, the consumption behaviour of individuals can change depending on external events that condition their access to and behaviour towards the goods they usually purchase. Faced with this, e-commerce offers the opportunity to expand the current offer and create new lines of business, and with these experiences it is expected to generate trust and comfort, which generates the emergence of new habits and behaviours that are expected to persist after the pandemic, therefore, digital platforms and channels are getting stronger day by day, where companies must take advantage of and react to the opportunities under the current situation.

In the new context the traditional channel and modern channel have been confronted in the pandemic, while the traditional channel markets products through points of sale in markets, neighborhood shops, kiosks, etc., where the customer had direct interaction with the seller and its logistics or planning has no presence, it was observed that due to the pandemic the population came closer to these points of sale to avoid contagions, therefore it had growth, despite having informality as a characteristic. For Peru Retail (2018), two years ago it was considered that "In the interior of Peru, the bodegas, markets and street markets continue at the top, reflecting more than 90% of sales; and in the case of the capital, 70%, which had turned the traditional channel into a strategic medium that allows participation, diversification and profitability". By 2020 and with the confinement it seems that things have changed drastically and although the traditional channels have fallen to opt for modern ones, there is still the possibility of rethinking their entry strategies and positioning to a new market.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In recent years, the Traditional Channel has faced a competitor that is advancing at a rapid pace: the Modern Channel. The various shop formats that make up this channel have been understanding new consumer habits, have improved the shopping experience, and have diversified their value offer. The use of technology has been one of their main tools, not only to provide better customer service but also to perform business intelligence. The fast competitor seems to have it all. However, the pandemic has brought changes in the game that were not seen coming and that in some way favor the Traditional Channel or give it a breath of fresh air to improve its value offer and its role vis-à-vis consumers. (Fundes Latinoamerica, 2020)

On the other hand, according to the study by Consumer 360 and DATUM International (2020), 3% of consumers still use a traditional channel to make their purchases and transactions, while 28% use modern channels. An evaluation of security levels at the entrance to establishments has also been carried out, with encouraging results, for example, in commerce, security levels of 87% were achieved, but within this level, 70% use the traditional channel, while 96% use the modern channel. On the other hand, an interesting fact to mention is the NPS (Net Promoter Score) Index, which measures the willingness of customers to recommend a brand, where establishments in the traditional channel were the least recommended, reaching -37%, a situation that will be difficult to recover.

Trends

As countries have progressed towards a 'new normal', many behavioral patterns have changed, such as the use of online food delivery (Zwanka & Buff, 2020, pp. 58-67). This situation was later, as at the beginning of COVID-19, the first reaction was to access health supplies, but as the population met this need, another important need was food, but to ensure its consumption with non-perishable products. We could conclude that the changes in patterns have an impact not only in the pandemic but also post-pandemic.

Changes in consumption habits are significant at 77%, which does not imply an increase in their consumption, but their satisfaction levels are relatively high at 56%, however, 45% still had some level of difficulty in making their purchases.

As we have seen, the purchasing processes have increased significantly and a usual form of payment is via internet bank transfers (40%), in this sense, the levels of security in the management of these systems must consider the necessary levels of security to generate benefits for consumers.

The process of implementing technologies that may still have a risk in terms of their use, could generate a greater risk if they are not used. In this sense, isolation presents different situations in the business, educational, social, etc. spheres, generating different dynamics in consumers.

Technology allows to monitor the global situation and make decisions of different kinds, on the other hand, it is necessary to rely on existing technology and be alert in the development or improvement of existing ones, being friendly and agile turning the virtual into a strategy of advancement and development, there is empirical evidence from studies tracking consumers' online search movement behaviours, i.e. search paths and patterns, from a need perspective that can contribute to a deeper understanding of how consumer need states lead to different online search behaviours and thus can provide marketers with a benchmark evaluation model.

It is necessary to investigate the processes through which consumer behaviour has had significant variations in the relationship with social networks and websites, generating levels of dependence that are expected to be maintained over time due to the agility of use, the reduction of time in the development of actions, which could eventually become irreversible due to the benefits.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

SOLUTIONS AND RECOMMENDATIONS

It is evident that the levels of adaptation to the use of ICTs for consumption have been effective, due to the development of user-friendly platforms, although there are some limitations due to the knowledge of age-related mechanisms and the mastery of ICTs, nevertheless, we are facing a new future, with significant changes in lifestyles, therefore, the levels of trust must be strengthened to guarantee the connection between consumers and companies.

After having developed the study, we can point out that we are living in different moments where society is developing accelerated processes in the adoption of technology, and society demands greater learning and knowledge generation in the new global context where the usefulness of the use of information and communication technologies (ICTs) in consumption must be rethought. It is possible to test prospective analyses to investigate the impact of new and different supply and demand channels and strategies differentiating goods of different priority in consumption, and it is also necessary to investigate the effect they have in coherence with cultural and anthropological aspects, particularly in rural areas.

FUTURE RESEARCH DIRECTIONS

It is also known that the marketing channels for food products currently differ significantly from those for health products, and it is therefore necessary to investigate whether these differences will continue and whether there is a trend towards new ways of converging on the same channel. It is also necessary to investigate to what extent the process of learning about the use of technologies could have an effect on the purchasing decision process and, in addition, to investigate the development of e-commerce and its influence on companies and customers in the post-pandemic situation, as well as to investigate intercultural applications generated on the internet, another topic of interest is related to written language through e-mail and social networks. Another important topic is the functioning of educational modalities, related to the cost-time and cost-benefit binomials.

We also consider that the differences in the digital divide may be more noticeable when comparing consumers in urban and rural areas, and that the information gathered will provide us with important elements for decision-making on the use of technology in different areas or sectors. Finally, to rescue and propose research perspectives related to ICTs, age groups, marketing channels in critical health situations (pandemics), among others.

CONCLUSION

In times of pandemic, ICTs have had a differentiated influence on the decisions of consumers of different age groups, generating challenges and opportunities in the process of adaptation of the different generations. In relation to age segments, some have had dominance and perhaps a greater advantage in the process of accessing digital environments, while others have had limited progress, we can mention that all of them have had active participation in the use of technologies to varying degrees.

The use of ICTs for e-commerce in Arequipa is progressively gaining ground in the Centennials and Millennials age segments, with a slight increase in Generation X and Baby Boomers, and it is possible that the trend will continue to grow even after the end of the pandemic. With different expertise, the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

different age segments analyzed already showed a growing trend in the use of ICTs before the pandemic, with less intensity in the baby boomers, and in this sense, we can speak of a greater dominance of consumers of younger generations.

Mistrust is an element mentioned in all age segments, and the greater the distrust, the more consumers prefer to pay in cash on delivery. This aspect is a challenge to further consolidate the progress of e-commerce, because, when asked about trust in social networks to make purchases, only 7% of digital natives stated that they trust them, while 74% of them only consider them to be moderate to not very trustworthy. Among digital immigrants, trust is only reflected in 1.3%. These results would indicate that there is a need to implement strategies to consolidate the issue of trust in digital platforms for all actors involved in their use.

The anthropological aspects of the use of electronic devices show that digital environments have been revealed as the necessary complement to speed up these processes, making possible a close relationship between consumers and economic activities, which over time have massified, thereby strengthening electronic commerce, a situation that will have an irreversible character, where the conditions of connectivity can strengthen the digital market. It has also been observed that, in underdeveloped societies, the cultural factor has had an important influence on the access and use of ICTs, due to the fact that, on the one hand, the link with these tools in these societies is later than in developed societies and, on the other hand, because they are more heterogeneous societies in terms of their cultural and social composition, thus delaying the massification of their use and therefore the adaptation and assimilation of the population for their use.

REFERENCES

Akerlof, G. A., & Shiller, R. J. (2015). *The economics of manipulation. How we fall like dupes into the plots of the market*. Planeta Colombiana S.A.

Baena Rojas, J. J., Cano Arenas, J. A., Jarrin Quintero, J. A., & Pérez Arroyave, H. R. (2014). Use of information and communication technologies for international negotiation: An advantage for Colombian companies? *Revista Ciencias Estratégicas, 22*(32), 279-294. Retrieved from https://www.redalyc.org/pdf/1513/151339264007.pdf

Ballesteros Riveros, D. P., & Ballesteros Silva, P. P. (2004). Competitive logistics and supply chain management. *Sciences et Techniques (Paris)*, 201–206. https://www.redalyc.org/pdf/849/84912053030.pdf

BBVA Research. (28 April 2016). *Peru: The retail sector.* Retrieved from https://www.bbvaresearch.com/publicaciones/peru-el-sector-retail/

Caceres Cabana, Y. D., Gonzalez, S., & Rivera Medina, J. (2020). Alpaca meat: constraints and reinforcements in consumption. In Anthropological approaches to understanding consumption patterns and consumer behaviour. IGI Global. doi:10.4018/978-1-7998-3115-0.ch016

CAF Development Bank of Latin America. (2020). The state of Latin America's digitalization in the face of the COVID-19 pandemic. *CAF Observatory of the Digital Ecosystem, 19*(3), 17. Retrieved from https://scioteca.caf.com/handle/123456789/1540

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CAF Development Bank of Latin America. (2020). *The state of Latin America's digitalization in the face of the COVID-19 pandemic* (Vol. 19). Retrieved from Observatorio CAF del Ecosistema Digital: https://scioteca.caf.com/handle/123456789/1540

Casco, A. R. (2020). Effects of the COVID-19 pandemic on consumer behavior. *INNOVARE Journal of Science and Technology, 9*(2).

Central Reserve Bank of Peru. (2016). *Informe Económico y Social Región Arequipa.* Retrieved from https://www.bcrp.gob.pe/docs/Proyeccion-Institucional/Encuentros-Regionales/2016/arequipa/ies-arequipa-2016.pdf

Consumer 360 and DATUM International. (2020). *Customer Experience and COVID-19 Comercio, Bancos y Farmacias protocols.* Retrieved from https://www.datum.com.pe/new_web_files/files/pdf/2020%20Experiencia%20del%20Cliente%20y%20Protocolos%20Covid-19.pdf

Cullen, R. (2001). Tackling the digital divide. *Online Information Review*, *25*(5), 311–320. doi:10.1108/14684520110410517

Diaz-Sarmiento, C., López-Lambraño, M., & Roncallo-Lafont, L. (2017). Understanding generations: a review of the concept, classification and distinguishing characteristics of baby boomers, X and millennials. *Revista Clío América*, 188-204. doi:10.21676/23897848.2440

Food and Agriculture Organization of the United Nations. (2017). *Food and Agriculture. Actions to advance the agenda of the 2030 Agenda and Sustainable Development Goals.* FAO. Retrieved from http://www.fao.org/3/a-i7454s.pdf

Fundes Latinoamerica. (2020). *Why are FMCG companies focusing - today more than ever - on the traditional channel?* Retrieved from https://www.fundes.org/por-que-las-empresas-de-consumo-masivo-se-estan-enfocando-hoy-mas-que-nunca-en-el-canal-tradicional/

Garcia, A. M., & Gil-Saura, I. (2016). *Innovating in retail trade: the influence of ICT and its effects on customer satisfaction*. doi:10.5295/cdg.10556am

Hao, N., Wang, H. H., & Zhou, Q. (2020). The impact of online grocery shopping on stockpile behavior in Covid-19. China agricultural economic review. *Revisión económica agrícola de China, 12*(3), 459-470. doi:10.1108/CAER-04-2020-0064

Harari, Y. H. (2020). *The world after the coronavirus.* Retrieved from http://virtual.iuc.edu.co/plugin-file.php/3512/mod_resource/content/1/2.%20FILOSOF%C3%8DA%20MAYO%2011%20-%2015.pdf

INEI. (2014). *Results of the Micro and Small Enterprise Survey 2013.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1139/libro.pdf

INEI. (2017). *National Census of Food Markets 2016.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1448/libro.pdf

INEI. (2018). *Technical report Estadísticas de las Tecnologías de Información y Comunicación en los Hogares.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/boletines/04-informe-tecnico-n04_tecnologias-de-informacion-jul-ago-set2018.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Internet World Sats. (2020). *Internet World Sats*. Retrieved from https://www.internetworldstats.com/stats2.htm

Jones, C., Alderete, M. V., & Motta, J. (2013). E-commerce adoption in Micro, Small and Medium-sized commercial and service enterprises in Cordoba. *Cuadernos de Administración, 29*(50), 164-175. Retrieved from https://www.redalyc.org/pdf/2250/225029797006.pdf

Junta de Andalucía. (2020). *Guide Using ICTs to reach new market niches and consumers.* Retrieved from www.formacion.andaluciaesdigital.es: https://www.formacion.andaluciaesdigital.es/catalogo-cursos/-/acciones/ficha/7809

Kantar. (2020). *A new world: Changes and prospects for shopping channels.* Retrieved from http://mkt.kantarworldpanel.com: http://mkt.kantarworldpanel.com/global/LATAM/Retail/Retail_COVID-19_v3.pdf

Kirk, C. P., & Rifkin, L. S. (2020, September). I'll trade you diamonds for toilet paper: Consumer reacts, copes, and adapts behaviors in the COVID-19 pandemic. *Journal of Business Research*, *117*, 124–131. doi:10.1016/j.jbusres.2020.05.028 PMID:32834208

Malone, A., Caceres Cabana, Y. D., & Taya Zegarra, A. (2020). Informal food systems and differential mobility during the COVID-19 pandemic in Arequipa, Perú. *The Town Planning Review*. Advance online publication. doi:10.3828/tpr.2020.61

Marcus, G. E., & Fisher, M. M. (1999). *Anthropology as Cultural Critique: An experimental moment in the human sciences.* doi:10.7208/chicago/9780226229539.001.0001

Martelo, R. J., Jimenez-Pitre, I., & Moncaris Gonzalez, L. (2017). Guía Metodológica para el Mejoramiento del Desarrollo de Software a través de la Aplicación de la Técnica Árboles de Problemas. *Información Tecnológica*, *28*(3), 87–94. doi:10.4067/S0718-07642017000300010

Martinez Coral, P. (2017). "Seguro mató a confianza": Challenges for the adoption of digital government in Colombia. *Academic Journals: Inclusion and Development*, *5*(1), 63–72. doi:10.26620/uniminuto.inclusion.5.1.2018.63-72

Martinez Dominguez, M. (2018). Access to and use of information and communication technologies in Mexico: determinants. *PAAKAT: Journal of Technology and Society, 8*(14). Retrieved from . doi:10.32870/Pk.a8n14.316

Masicampo, E. J., & Baumeister, R. F. (2011). Consider it done! Plan making can eliminate the cognitive effects of unfulfilled goals. *Revista de personalidad y psicología social, 101*(4), 667-683. doi:10.1037/a0024192

Mullainathan, S., & Shafir, E. (2016). *Scarcity: Why does having little mean so much?* Fondo de Cultura Económica.

Nielsen. (2020). *COVID-19 will affect low-income consumers in Latin America the most.* Retrieved from https://www.nielsen.com/pe/es/insights/article/2020/covid-19-afectara-mas-a-los-consumidores-de-bajos-ingresos-en-latinoamerica/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ogbum, W. F., & Thomas, D. S. (1922, May 8). The influence of the business cycle on certain social conditions. *Journal of the American Statistical Association*, *18*(139), 324–340. doi:10.1080/01621459 .1922.10502475

Ortega-Vivanco, M. (2020, March). Effects of Covid-19 on consumer behaviour: The case of Ecuador. *RETOS. Journal of Management Science and Economics*, *10*(20). Advance online publication. doi:10.17163/ret.n20.2020.03

Pérez-Escoda, A., Castro-Subizarreta, A., & Fandos-Igado, M. (2016). The digital competence of Generation Z: keys for its curricular introduction in Primary School. *Scientific Journal of Educommunication, 49*, 71-80. Retrieved from https://www.redalyc.org/pdf/158/15847434008.pdf

Peru Retail. (2018). *Peru: Traditional vs. modern channel.* Lima. Retrieved from https://www.peru-retail.com/peru-canal-tradicional-vs-canal-moderno/

Peru Retail. (2020) *What are the most visited shopping channels by Peruvians before Covid-19?* Retrieved from https://www.peru-retail.com/cuales-son-los-canales-de-compra-mas-visitados-por-los-peruanos-ante-el-covid-19/

Phillip Kottak, C. (2011). Cultural anthropology. McG.

Regional Government of Arequipa. (2020). *Regional Executive Resolution N° 192-2020-GRA/GR.* Retrieved from www.regionarequipa.gob.pe: https://www.regionarequipa.gob.pe/Cms_Data/Contents/GobRegionalArequipaInv/Media/Resolucion.Detalle/2020/R.E.R/RER-192-2020.pdf

Rivoir, A., Morales, M. J., & Casamayou, A. (July 18, 2019). Uses and perceptions of digital technologies in older people. Limitations and benefits for their quality of life. *Austral Journal of Social Sciences*, 295-313. doi:10.4206/rev.austral.cienc.soc.2019.n36-15

Serrano, C. L., Peña Lapeira, C. J., & Laverde Guzmán, M. Y. (2018). Influence of ICT in the economic development of Colombia. *Revista Ciencias de la Información*, 3-10. Retrieved from http://cinfo.idict.cu/index.php/cinfo/article/view/854

Sheth, J. (2020). Impact of Covid-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research 117*, 280-283. doi:10.1016/j.jbusres.2020.05.059

Superintencia de Banca. Seguros y AFP - BID. (2012). *Pilot Survey on Access to and Use of Financial Services 2012.* Retrieved from https://www.sbs.gob.pe/Portals/0/jer/est_incl_finan/Informe%20de%20resultados%20inferenciales_.pdf

Telecom Advisory Services, LLC. (2015). *The Digital Ecosystem and Economy in Latin America*. Retrieved from http://www.teleadvs.com/wp-content/uploads/Presentacion_IGF_v3.pdf

Thomas, F., Handdon, L., Gilligan, R., Heinzmann, P., & De Gournay, C. (2005). Cultural Factors Shaping the Experience of ICTs: An Exploratory Review. *ResearchGate*. Obtenido de https://www.researchgate.net/publication/251783055_Cultural_Factors_Shaping_the_Experience_of_ICTs_An_Exploratory_Review

Varian, H. R. (2014). *Intermediate Microeconomics A Modern Approach* (9th ed.). Theresia Kowara.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Weinmann, M., Schneider, C., & Vom Brocke, J. (2016). Digital Nudging. *Business & Information Systems Engineering*, *433-436*. Advance online publication. doi:10.2139srn.2708250

World Economic Forum. (2019). *What are ICTs?* Retrieved from en.weforum.org: https://es.weforum.org/agenda/2019/02/que-son-las-tics/

Zwanka, R., & Buff, C. (2020). Generation COVID-19: A conceptual framework of consumer behavioral changes to be caused by the COVID-19 pandemic. *Journal of International Consumer Marketing*, *33*(1), 58–67. doi:10.1080/08961530.2020.1771646

KEY TERMS AND DEFINITIONS

Arequipeño: Gentilicio of the people born in Arequipa.

Consumer: People of both sexes, from different age segments who develop various ways to purchase goods and services from providers in the context of the Covid-19 pandemic, through the use of information technologies.

Consumer Trends: Evolution of the changes and strategies consumers have chosen to purchase goods and services as a result of the COVID-19 pandemic.

Digital Technologies: Information technology tools used by consumers in email services, web usage, data storage, e-banking, virtual communities managed through devices such as computers, mobile phones and tablets.

COVID-19: Pandemic that has affected the health of the population with effects on other socioeconomic components in the context of the city of Arequipa.

Generational Groups: Groups of individuals who purchase goods and services, classified by age criteria in accordance with the definitions of Baby Boomers, Generation X, Y and Z.

Informal Sector: Persons or organizations that carry out precarious economic activities with the aim of self-employment, but without having fulfilled legal requirements of formality.

Marketing Channels: Different routes that goods and services take, managed by economic actors from production processes to delivery to consumers in the pre-pandemic and pandemic stages.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 16
Artificial Intelligence in Marketing:
A Review of Consumer–AI Interactions

Özge Sığırcı
Kırklareli University, Turkey

ABSTRACT

The purpose of this chapter is to shed light on the consumer-AI interaction in the marketplace. By this aim, the chapter uses a literature review approach. The previous literature examining AI from a consumer behavior perspective is reviewed, and the findings are compiled in a meaningful flow. According to the review, we see that the traditional marketplace is shaped by AI from only human-to-human interactions to human-to-AI and AI-to-AI interactions. In this new marketplace, while consumers interact with AI, they gain new experiences and feel positive or negative because of these experiences. Also, they build different relationships with AI, such as servant, master, or partner. Besides these relationships, there are still concerns about AI that are related to privacy, algorithmic biases, consumer vulnerability, unemployment, and ethical decision making.

INTRODUCTION

The artificial intelligence (AI) concept was born in the mid-twentieth century (Feng et al., 2020) and can be defined as the "computational agents that act intelligently" (Poole & Mackworth, 2010, p. 3). Human intelligence is imitated by machines in AI. Thus, AI can complete the tasks such as visual perception, speech recognition, and decision-making under uncertainty that normally requires human intelligence (Gillath, et al., 2021; Rossi, 2018; Russell & Norvig, 2016).

Today AI is used by companies and it is in the life of consumers. For example, consumers can talk to a chatbot when they call a company or interact with the robot employees working at hotel receptions when making a hotel reservation. Product recommendations in online shopping are made by an AI system even if the consumer is not aware of. Besides, there are smart home appliances such as smart refrigerators in today's homes that can talk to each other to find out what is needed at home and make the orders

DOI: 10.4018/978-1-7998-6985-6.ch016

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

on behalf of the consumer. There are robot vacuum cleaners like iRobot which can do the cleaning for its owner. There are digital assistants such as Alexa or Siri that wait to hear their names to satisfy their owners' requests changing from learning today's weather to making a restaurant reservation. Also, AI-enabled self-driving cars is just around the corner. As can be seen from all these examples, AI is in the middle of consumers' life both with the companies' marketing activities and the products or services consumers directly use. Market growth rates also tell us that AI will penetrate more into the consumers' life in the near future (Bughin et al., 2017).

At this point, questions about the state of consumer-AI interaction comes to mind. What does AI mean to consumers, does it satisfy consumers to be served by an AI agent instead of a human, what sort of experiences and relationships do consumers have with AI, in which situations do consumers resist AI and what are consumers' concerns about AI? The answers to these questions are critical for companies. Because the implementation of AI in marketing has changed the way companies interact with their consumers and offered companies many benefits. For instance, with the help of AI technologies, companies can manage consumer data more efficiently and better understand their consumers and their shopping patterns, which can positively affect consumer satisfaction (Ameen et al., 2020; Evans, 2019). But AI system investments that are not wanted and accepted by consumers can create problems such as lower consumer satisfaction, negative consumer perception and behavior. Understanding how consumers perceive AI, the types of relationships they have with AI and their concerns toward AI will help companies design right AI-enabled products and services, take precautions for consumer concerns and reach their target customers with the right marketing communication strategies. Thus, the main aim of this chapter is to shed light on the consumer-AI interaction in the marketplace. By this aim, we reviewed the previous literature examining AI from a consumer behavior perspective and compiled the findings in a meaningful flow under this chapter.

This chapter could be beneficial for different parties such as practitioners, researchers, and students. We hope the practitioners can use this chapter to understand the consumers' perception of AI better and use these findings to design their AI technologies better. The researchers can see the previous research, detect the gaps in the literature, and design future research questions. Also, the students can learn the dynamics of AI in businesses and consumer behavior.

The chapter begins with introducing AI and its foundations and examining the usages of AI in marketing by giving real-world examples from a variety of sectors. Then we focus on how AI has transformed the service encounter, how consumers interact, have experiences, build relationships with AI and consumers' attitudes toward AI. It is followed by the concerns both consumers, researchers, and critics have about AI. In the final part of the chapter, we offer a conclusion, based on the discussion of reviewed literature, and directions for future research.

What is Artificial Intelligence

Artificial Intelligence is about computer systems that mimic the intelligence of humans. So, the question is, what do we mean by mimicking human intelligence? We mean that these computer systems have the ability to perceive the world around them and take actions based on these perceptions (Overgoor et al., 2019) and perform typical human mind functions such as perception, cognition, and conversation (Longoni, Bonezzi & Morewedge, 2019). Visual perception, speech recognition, decision-making, memory, reasoning, problem-solving, and goal-oriented actions are some of the abilities of AI that position it close to human intelligence (Paschen et al., 2019; Rossi, 2018).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Artificial Intelligence systems perceive and process the input data to take the aforementioned actions (Jarek & Mazurek, 2019). AI uses many tools to perform its actions, such as machine learning, deep learning, natural language processing, neural networks, rule-based expert systems, physical robots, and robotic process automation (Davenport & Ronanki, 2018). Among these technologies, machine learning has a special place. Machine learning is a subset of AI (Russo, 2020), and with the help of machine learning, computers can learn by themselves from the data by finding patterns in the data (Jarek & Mazurek, 2019). With the advancements in machine learning technology, computer systems not only follow the instructions of the program, but also they learn from the training data by recognizing the patterns in the data and developing algorithms based on these patterns. As a result, they perform various tasks, such as making predictions and recommendations based on their learning (Du & Xie, 2021; Russo, 2020). Examples of AI include personal helpers like Siri and Alexa, credit scoring assistants like IBM's Watson, humanoid robots that can be used as greeters or service providers like Pepper, and self-driving cars.

Foundations of Artificial Intelligence

The roots of artificial intelligence can be traced back to the 1940s (Haenlein & Kaplan, 2019). In 1942 Isaac Asimov published a short story named "Runaround," which was built on Asimov's "Three Laws of Robotics". The story has inspired scientists from a variety of fields for generations. At the similar time period, mathematician and computer scientist Alan Turing, who was working for the British government, invented a code-breaking machine to break the Enigma code used by the German army. Breaking of the Enigma code by a machine, led Turing to think about the intelligence of machines (Haenlein & Kaplan, 2019). In 1950, he pointed out artificial intelligence and proposed Turing Test to test the intelligence of the machines in his influential article "Computing Machinery and Intelligence" (Turing, 1950).

In 1956, John McCarthy and his colleagues organized the *Dartmouth Summer Research Project on Artificial Intelligence*. They invited researchers from various fields to discuss developing machines that simulate human intelligence (McCarthy et al., 1955). This workshop was considered the beginning of the AI field, where the name *AI* is officially used, and the founders of the AI field came together (Feng, et al., 2020; Haenlein & Kaplan, 2019; Ma & Sun, 2020). Since then, research trying to understand AI and the development of AI entities have continued and progressed. Today, image recognition, speech processing, and autonomous driving are some of the possible tasks performed by AI (Ma & Sun, 2020). With the help of technological advancements in Big Data and computing power (Haenlein & Kaplan, 2019), AI has impacted and transformed many fields from finance and education to healthcare and biology (Huang & Rust, 2018; Rust, 2020). Business and marketing are also one of the most benefitted fields from AI.

To better understand where we are today in the developmental stages of AI, a discussion of the AI types can be necessary. Literature mentions two types of AI: *weak (narrow) AI* and *strong AI*. *Weak (narrow) AI* is the type of AI that can mimic human intelligence by analyzing large amounts of data (Russell & Norvig, 2016). Though it is trained to solve a specific problem, it needs training for solving other problems (Wirth, 2018). It does not mean that weak AI cannot mimic human intelligence or perform tasks successfully, but it is comparatively weak against human intelligence. So, it can be said that, though weak AI systems can be very powerful in their specifically trained area, they are not as flexible as human intelligence in other areas (Wirth, 2018). On the other hand, *strong AI* is about the machines having a conscience (Russell & Norvig, 2016), and we can speak of context awareness (Davenport et al., 2020). In strong AI, we expect machines to think holistically, give context-specific responses, and as a result complete more complex tasks (Huang & Rust, 2018).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Considering the recent developments in AI, researchers think that it is not easy to categorize some of today's advanced systems as weak/narrow AI and offered the term '*hybrid AI'* for the systems that integrate multiple weak AI systems (Greenwald, 2011). Also, they offered the idea to conceptualize the levels of artificial intelligence as a continuum to position the systems such as Google's DeepMind AlphaGo or IBM's Watson at a higher level than positioning them as just weak/narrow AI (Davenport et al., 2020; Wirth, 2018). Though it should be noted that, while it is possible to position some of today's systems to an upper level than weak AI, AI systems today is still far from being at the level of strong AI, and reaching a context-awareness level looks like a low chance by 2050 (Davenport et al., 2020; Müller & Bostrom 2016).

Different Usages of AI in Marketing

When we look at the usage of AI in businesses, we see a broad set of applications in various sectors such as retailing, health care, public sector, hospitality, and education. This broad usage areas lead consumers to interact with and have experience with AI. Sometimes consumers are aware that they are interacting with an AI and sometimes they are not aware of that, even if there was an AI agent at the background. But in both cases a consumer experience occurs, and this experience affects consumers' satisfaction or dissatisfaction with the brand. Hence, here we will try to examine the different usages of AI in the marketing activities of businesses from different sectors, based on real-world examples. Examining the AI applications in businesses -especially in marketing and consumer behavior-related areas- led us to think that it can be appropriate to categorize these usages under three categories. First of all, there are AI systems that are kind of an employee for the firms and directly interact or communicate with the customers. In the second group, there are AI systems that help the firm with some of the tasks or perform the tasks on behalf of the firm but do not have direct interaction with the customers. In the final group, there are AI systems that are actually the product or the service itself the company offers.

*AI systems that direct customer-AI interaction occur*s: Under this category, we will examine AI systems that work as the employees of the company and directly interact with customers. The first way companies use AI under this category is as chatbots or virtual assistants. Companies use chatbots or virtual assistants to interact with their customers and answer their questions. Since these virtual assistants can provide 24/7 customer support and costs lower than human employees, many companies choose them as their employees (Pantano & Pizzi, 2020). For example, Ojo Labs, a real estate company, developed a chatbot to answer potential homebuyers' questions via messaging (Longoni & Cian, 2020; Wiggers, 2019). Also, another real estate company REX Real Estate is also using a voice/chatbot for home showings (Silverstein, 2018). An alcoholic drinks company, Diageo's chatbot called 'Simi-Your Personal Bartender' gives consumers cocktail recipes based on the consumers' preferences and the ingredients they have at home (Bhatia, 2016). IPsoft developed AI digital assistant Amelia in 2014, and many companies now use it as an AI employee to serve as a customer support agent (Hwang & Park, 2020). She has many human interaction elements such as conversation, expression, emotion, and understanding (Amelia, n.d.). Another example is Conversica AI virtual assistant; it is a chatbot that interacts with customers and potential customers and helps companies both in sales and marketing activities (Davenport et al., 2020).

Other usages of AI adopted by companies that have direct interaction with the customers can be in the payment or recommendation systems. For example, KFC in China uses a facial recognition-based AI system so that customers can buy their food and make their payments by only showing their faces to the system (Thales, 2018). Similarly, we see the usage of image recognition-based technologies in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the beauty and fashion sectors too. Revieve has an AI Skincare advisor that detects the consumers' skin and recommends suitable products for their skin after uploading a selfie (Revieve, n.d.). In the fashion industry, virtual mirrors help consumers see the collection on themselves virtually and recommend appropriate pieces to their appearance and taste (Jarek & Mazurek, 2019).

From the point of different sectors, we see the usage of AI in financial services. For example, Plum and Betterment are AI assistants that help people manage their money by saving, investing, or reducing their bills. These financial assistants analyze people's funds and expenses and develop customized savings plans, and they provide these financial services at a lower cost than a human financial expert. Moreover, Upstart is an AI lending service that uses non-conventional variables to calculate consumers' credit risk and help consumers' access to credit and offer lower loss rates (Upstart, n.d.).

Tourism and hospitality is another sector that uses AI. Although many companies use AI in purely virtual forms, such as virtual assistants, as we mentioned previously, AI can also be embedded in robots that have physical embodiment (Davenport, et al., 2020). Previous literature also shows that consumers prefer to interact with robots than with AI without a physical embodiment (Davenport, et al., 2020). We see examples of humanoid service robots in the tourism and hospitality sector in addition to virtual assistants. For example, KLM introduced Robot Spencer to serve in Schiphol Amsterdam Airport. Robot Spencer scans passengers' boarding passes and guides them to their departure gate; also, it adjusts its speed according to the passengers and informs them about the departure gate distance (KLM, 2016). KLM started using their new robot; KLM Care-E, a self-driving luggage trolley with voice and physical movement abilities to scan passengers' boarding passes, carry their luggage, and guide them to their gates (Klmcare, n.d.). In hotels, AI is used in the form of autonomous check-ins or room delivery robots (Tussyadiah, 2020). For example, Hotel Trio in Healdsburg, California, uses service robot Rosé to deliver wine to the rooms (Romano, 2020). A unique example is Henn na Hotel in Japan, whose staff are mostly robots. It is possible to see dinosaur robot receptionists or robots in the cloakroom (Henn na Hotel, n.d.). Other examples can be the robot barista called Tipsy Robot that works at a bar in Las Vegas, or the Café X, robotic coffee bars.

The last sector that we commonly see the AI applications, and we want to mention here is the healthcare sector. Although the usage of AI in the healthcare sector is manifold, under this category, we will only mention the AI systems that have direct interaction with the patients. For example, the UK's National Health Service (NHS) tried a medical chatbot in its non-emergency helpline to help patients give medical advice (Vincent, 2017). There are virtual nurses like Care Angel too.

AI systems that no direct customer-AI interaction occurs: Firms also use AI systems to help them manage some of their tasks, but consumers may not have a direct interaction with these AI systems. The first example can be the AI recommendation systems; many firms use AI for recommending customers new products or services based on their past purchases or rated items. Companies such as Amazon, Netflix, and Spotify use these AI-enabled recommendation systems to recommend consumers products or services they might like. Many beauty brands also make recommendations about their products by using AI (Longoni & Cian, 2020). We see the usage of AI recommendation systems in clothing and styling businesses as well, such as Stitch Fix and Trendy Butler (Davenport et al., 2020). AI-enabled recommendation systems significantly affect the companies' revenue, such that Amazon's recommendation system generated 35% of the company's revenue (Morgan, 2018).

Healthcare is another sector that uses AI. By offering expert-level accuracy, cost-effectiveness, and faster processes (Esteva et al. 2017; Gallagher, 2017; Leachman & Merlino, 2017), AI helps transform and develop healthcare services. AI is being used in healthcare to help physicians in tasks like evaluat-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing the radiological images, pathology slides, or EMR results of patients (Mintz & Brodie, 2019). For example, IBM's Watson is designed to help doctors diagnose heart diseases (Hinchliffe, 2017). Firms also use AI to understand online WOM (Netzer, 2012; Tirunillai & Tellis, 2014), to do stock management in warehouses like Amazon Kiva, to develop creative ad campaigns (Jarek & Mazurek, 2019), and to optimize prices (Antonio, 2018). With the help of AI, even online retailers will be able to transform their business models from shopping-then-shipping to shipping-then-shopping. This means that retailers will ship to the customer based on the accuracy of their predictions even before the customer orders the product (Davenport et al., 2020).

AI systems which are the product or the service itself that the company offers: Voice-controlled virtual assistants such as Apple's Siri, Google's Assistant, Amazon's Alexa are major examples for this category that have pervasive usage. They are the virtual helpers of the users, and the users do not have to learn any specific commands to interact with these assistants because these assistants listen for their names and respond to natural language (Jones, 2018). They can help with a variety of tasks; sending a text, setting a reminder, making restaurant reservations, playing your favorite music, giving a real-time answer about the weather or traffic are just a few examples we can mention about their capabilities. There are also robots used for homecare like iRobot, or the companion robots like Buddy, and the commercialization of autonomous vehicles is predicted to be inevitable (Lanctot, 2017). These assistants or autonomous agents are not only making life easier for their users, but also they are helpful for their manufacturers. By analyzing the data collected by these assistants, which are critical personal level data, brands can understand their consumers more deeply and take strategical and tactical decisions (Jones, 2018).

AI and New Marketplace

As can be understood from the examples given above, AI and usage of AI in marketing have transformed the traditional marketplace. In this section, we will examine how the marketplace service encounter has been transformed by AI. In traditional marketplace service encounters, we see that both parties of the encounter -the customer and the service provider- are human. In other words, we are used to seeing interhuman (human-to-human) service encounters in the marketplace. For instance, when we think about a hotel reception service in the traditional marketplace, both the receptionist and the customer are human. However, today's advances in technology reveal service encounters more often where the customer or service provider or both are artificial intelligence.

Robinson et al. (2020) categorized today's new service encounter as; (1) human customer-to-human frontline employees (interhuman) service encounters, (2) human customer-to-AI frontline employees service encounter, (3) AI customer-to-human frontline employees service encounter, and (4) AI customers-to-AI frontline employees service encounter.

1. Human Customers-to-Human Frontline Employees: This is the traditional interhuman service encounter, where both the customer and frontline employees are human.
2. Human Customers-to-AI Frontline Employees: This is an "interspecific" type of interaction where the customer is human, and the AI frontline employee is interacting with the human customers on behalf of the firm (Robinson et al. 2020). For example, a customer served by a robot receptionist at the robot hotel, Hotel Henn na in Japan, has this type of relationship. As a more frequent usage, a human customer interacting with a company's chatbot to make orders or get some information also experiences this type of relationship.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. AI Customers-to-Human Frontline Employees: On the other hand, we now see AI playing the role of the customer and interacting with a human service provider. A digital assistant calling a doctor's office on behalf of its owner to make an appointment and answered by a human employee can be an example of this encounter. These AI customers can mimic human voice so perfectly that they are even criticized for deceiving the human service exchange partner (Robinson et al. 2020).
4. AI Customers-to-AI Frontline Employees: We do not only see AI consumers interacting with human frontline employees, but also there can be an AI frontline employee at the other end of the service exchange. This service encounter is labeled as "inter AI" service encounters, where both the customer and the firm are represented by an AI agent (Robinson et al. 2020). In other words, AI agents communicate with each other on behalf of the customer and firm without any human involvement. This type of AI-to-AI communication occurs in smart service systems (Robinson et al. 2020). For example, Google Home can communicate with Amazon's Alexa and order food through FoodMonkey application which is also another AI-enabled system (Voice2Biz, n.d.)

To understand the broader interactions of human and artificial intelligence in the marketplace, it is possible to broaden the service encounter framework of Robinson et al. (2020) by including the interactions within customers and interactions within employees. For instance, in this new service encounter framework, one can also think about the interactions between human customers and the AI-customers (e.g., voice controlled smart assistants) and the interactions between human frontline employees and AI-frontline employees too.

The categorization and examination of today's service encounter shows us that, with the inclusion of AI systems, today's service encounter has turned into a new place. It is critical for marketers to understand consumers' perception, attitudes, relationships and behaviors in this new marketplace. Moreover it is critical to understand who the new AI consumer is. While MacInnis and Folkes (2010) examined the boundaries of consumer behavior research, they underlined the role of "people" playing the consumer role. However, today, we see that objects with AI also start playing the role of consumer. Thus, in the light of Hoffman and Novak (2018: 1198), who considered smart objects "as consumers with their own consumption experiences", and proposed assemblage theory framework which "strongly implies that smart objects play a role in consumption-related processes", it is maybe time to question the boundaries of who the consumer is and the boundaries of consumer behavior research in that sense. Hence, it is time for firms to think about adapting their marketing strategies for these AI consumers and researchers to focus their research effort on

better understanding AI consumers. In relation with this new service encounter perspective, in the following sections, we will focus on the previous research findings to explain the consumers' experiences and relationships with AI, their attitudes toward and concerns about AI.

Consumer Experiences and Consumer Relationships with AI

Since there is a consumer experience in every interaction through the customer journey (Lemon & Verhoef, 2016), consumers' interactions with AI in the marketplace also create consumer experiences. From the consumer point of view, Puntoni et al. (2021) classified these experiences under four groups: data capture experiences, classification experiences, delegation experiences, and social experiences.

Data Capture Experience: Data capture experience is the experience of consumers' sharing their data with AI. Consumers can share their data with AI in different ways; they can either intentionally

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

transfer their data to AI to use the services AI offers or the AI can collect consumers' data from their "footsteps" they leave behind (Puntoni et al., 2021). For instance, consumers can intentionally share their specific preferences with a digital assistant to get customized recommendations. As a consequence of data capture experience consumers can either feel served by the AI or exploited (Puntoni et al., 2021).

Classification Experience: Classification experience is the consumers' perception that they are getting personalized predictions because they are labeled as consumers with certain types of tastes. As the consequences of classification experiences, consumers might feel understood or misunderstood (Puntoni et al., 2021).

Delegation Experience: Delegation experience is the experience of consumers when they delegate some of their tasks to the AI and enjoy the AI solution. For instance, consumers asking their digital assistants to play their favorite Spotify list or make an appointment at the doctor's office are having a delegation experience. Consumers can feel empowered by delegating some of their tasks but can also feel replaced when too many tasks can be handled by AI (Puntoni et al., 2021).

Social Experience: The last experience is the social experience. Since AI has the capability to communicate reciprocally, consumers who are interacting with AI have a social experience. Sometimes, consumers know that their interaction partner is an AI agent, but some of the other times, they may not realize it. The consequences of social experience can be feeling connected on the positive side, but feeling alienated on the negative side (Puntoni et al., 2021).

Consumers interact with AI and have either positive or negative experiences with them and form relationships with them. Thus, it is possible to question the types of relationships consumers might have with AI. In the marketplace, AI plays the role of customer and frontline employee. Besides, AI is also in consumers' lives in other forms such as digital assistants. Consumers interact with these digital assistants; they ask questions and have conversations with these digital assistants. On the other hand, firms imbue human-like features such as voice and names to make consumers adopt these assistants and build strong relationships with them (Aggarwal & McGill, 2007). All these interactions with the AI bring the questions of what type of relationship does AI and consumers have and how consumers define their relationships with the AI. In an effort to understand the type of relationships consumers build with AI, Schweitzer et al. (2019) conducted in-depth interviews with the users of voice-controlled smart assistants. They found that the relationship consumers build with their digital assistants is a relationship that is similar to human relationships. More specifically, consumers have servant, partner, and master relationships with their digital assistants.

Artificial Intelligence as a Servant: Some of the consumers perceive AI-based assistants as servants because they are dependent on the user's orders and commands and help the user achieve their tasks. In this relationship, while the user perceives the AI-based assistant as the servant, he/she perceives him/herself as the master. In this relationship type, consumers perceive the assistant as beneficial for their life, as helpful for them to extend their possibilities, and as a part of their self (Schweitzer et al., 2019).

Artificial Intelligence as a Master: Consumers who picture AI-based assistants as a master have concerns about whether the information that comes from these assistants could be trusted and whether the next move of them could be predictable. Consumers think that everything is up to the goodwill of the digital assistant and see themselves as slaves that has to obey the rules of the digital assistant. Consumers who generally have more negative experiences with these digital assistants and who could not manage to achieve their tasks with the help of the digital assistants generally perceive their relationship with them as the digital assistant as the master. For instance, situations that the digital assistants cannot understand the command of the consumer and answers unrelatedly reinforce the idea that the digital assistant is the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

master with its own autonomy, and they can never be trusted. Consequently, consumers who perceive the AI-based assistant as a master resist the relationship with the AI (Schweitzer et al., 2019).

Artificial Intelligence as a Partner: Also, there are a group of consumers who perceive the AI-based assistants as an equivalent partner. In this relationship type, AI is perceived to have its own life and character. Consumers tend to ascribe more human-featured characteristics to the digital assistant in this relationship. In this type of relationship, consumers are generally willing to treat the AI-based assistant as a person, invest time in their relationship with it and help him/her learn from them and grow (Schweitzer et al., 2019). But these investments to the relationship can turn into a disappointment easily when consumers feel like they are not getting enough answers and progress from the AI-based digital assistant (Schweitzer et al., 2019).

Consumers' Attitudes Toward AI

Although artificial intelligence can outperform human judgments in many domains and artificial intelligence offers superior accuracy than human intuition, consumers can still resist artificial intelligence. Researchers call people's aversion to algorithms and preference of humans over algorithms in a variety of decision tasks as "algorithm aversion" (Dietvorst, Simmons & Massey 2015). According to previous literature, the reasons behind consumers' resistance to artificial intelligence (algorithmic models) are numerous. Consumers' beliefs about artificial intelligence, their individual characteristics, or the context of consumption can affect consumers' resistance to AI. Here, we will mention the findings of the previous literature on factors affecting consumers' attitudes toward AI.

Individual factors or consumer characteristics have an effect on consumers' opposition toward artificial intelligence. It is shown that opposition to artificial intelligence is higher among individuals with lower scientific knowledge, higher distrust to the appropriate functioning of science and technology, less prone to innovation and change, and expressing egalitarian values (Lobera, Fernández Rodríguez & Torres-Albero, 2020). Also, women's perception of risk is higher in general (Gustafsod, 1998), and they are willing to take less risk (Byrnes, Miller & Schafer, 1999). Because of their higher risk perceptions, women are more likely than men to resist artificial intelligence, especially when outcomes are perceived as risky (Castelo & Ward 2016). Consumer identity issues are also shown to influence consumers' adoption of AI. Consumers want to take the credit for their own abilities for the positive outcomes when it comes to tasks that they perceive as central to their identity, such as cooking (Leung, Paolacci, & Puntoni 2018). Thus, consumers may resist AI for the tasks they find central to their identity (Castelo, 2019).

Consumption context is another critical point determining whether consumers prefer artificial intelligence or resist it. Consumers prefer human labor to artificial intelligence labor for products, services, or consumption contexts where symbolic value is higher. Because consumers have stronger uniqueness motives in symbolic consumption contexts and perceive that human labor can satisfy their uniqueness motives more than artificial intelligence labor, they have a tendency not to prefer artificial intelligence in symbolic consumption contexts (Granulo, Fuchs & Puntoni, 2020). For similar reasons, consumers resist artificial intelligence in comparison to human labor in medical contexts too. Consumers believe that their characteristics, symptoms, and circumstances are distinct from an average person and unique. They think that, an artificial intelligence service provider would not consider this uniqueness as much as a human service provider. Because of these concerns about uniqueness neglect, they do not prefer artificial intelligence providers in medical contexts (Longoni, Bonezzi & Morewedge, 2019).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Characteristics of the tasks and consequences of the tasks also play essential roles in consumers' resistance or adoption of artificial intelligence. From the point of task characteristics, consumers generally resist AI for tasks that are subjective and involve intuition and affect (Castelo, 2019) since they perceive that AI does not have enough affective capabilities to perform such tasks (Castelo et al. 2018). Also, tasks can differ in their consequences, such that recommending a book is less consequential, but diagnosis in medical domains can be more consequential (Davenport et al., 2020). Consumers resist AI when the tasks are more consequential, and the perceived risks are higher (Castelo et al. 2018; Castelo & Ward 2016). Similarly, consumers' beliefs about AI, such as the lower empathy capabilities, influence consumers' resistance to AI too. In their research, Luo et al. (2019) showed that since consumers perceive AI as less empathetic, the purchase rates decreased by 75% when consumers are told that their interaction partner is an AI bot.

On the other hand, literature shows us that, in some situations, consumers do not resist artificial intelligence and, moreover, prefer artificial intelligence over human intelligence. One of the researches trying to understand under which conditions consumers prefer artificial intelligence over human intelligence underlines the impact of product/service type. Longoni and Cian (2020) investigated consumers' preference to get recommendations either from an artificial intelligence agent or from a human. They showed that consumers prefer artificial intelligence-based recommendations over human-based recommendations when the utilitarian attributes are more salient than hedonic attributes. Consumers believe that artificial intelligence recommenders use cognitive evaluative dimensions such as facts, rationality, and logic but human recommenders use affective evaluative dimensions such as sensory experiences, emotions, and intuition. Thus, they believe that artificial intelligence-based recommenders are more competent than human recommenders, where utilitarian attributes are more important than hedonic attributes and prefer getting recommendations from an artificial intelligence in the utilitarian realm (Longoni & Cian, 2020).

Nevertheless, not resisting AI comes with its own costs. In occasions where consumers do not resist AI, they might tend to attribute more responsibility to AI than human service providers. Hong (2020) compared people's responsibility attribution to AI vs. human driver in case of an accident and found that the AI driver is blamed more than a human driver in case of an accident for the injury or death of the victim, and this blame attribution was even higher when the consequences of the accident are more severe.

Concerns About AI

Although there are many benefits of AI both for consumers and companies, there are concerns about AI from different groups of people such as consumers, society, and researchers. Understanding these concerns might help researchers direct their future research into these areas, help businesses answer these concerns by their communication strategies and help manufacturers design their AI products and services accordingly. Thus, here we will mention the main concerns about AI pointed out in the literature.

Privacy and Cybersecurity Concerns

The data-centric nature of AI products and services offers many benefits to consumers, such as geolocalization, health monitoring, digital assisting, recommending products or services, but also causes harms to consumers. Although there are beneficial services for consumers in return, firms' easy access to consumer data with the help of AI products has brought concerns about consumer privacy and cybersecurity both from consumers and critics. These consumers concern about the privacy and the security

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of their data are so salient that they can even decide to restrict their online transactions (Akhter, 2014) and oppose artificial intelligence because of their negative perceptions about the treatment and privacy of their personal data (Lobera, Fernández Rodríguez & Torres-Albero, 2020).

The source of consumer concerns about privacy comes from the firms' usage of their data in ways that violate their privacy (Hwang & Park, 2020). On the other hand, privacy can be defined as a person's right to have a control on the information about him/herself (DesJardins, 2014). So, the question is, when privacy violations occur? The main element of the answer to this question is the informed and voluntary consent of the person. We can talk about privacy violations "whenever personal information is collected or used without the informed and voluntary consent of the person" (Du & Xie, 2021, p. 7). Thus, the AI products -especially those that are high on interactivity, create more consumer privacy concerns. Because these types of products not only collect voluminous data but also the variety of the data they collect are diverse (Du & Xie, 2021). For instance, as an AI-enabled product and high on interactivity, the Apple watch, both collect a large quantity of data and a large variety of data from textual, visual, audio to verbal and sensory (Du & Xie, 2021). But, consumers may not be even aware of that much data collection which can lead to consumer data privacy concerns. In sum, firms can collect and use a variety of consumer data such that; they can collect and use personal information to develop the performance of AI, they can collect and use consumption behavior information to predict what would consumers buy in the future and to recommend appropriate products or services to them and they can collect and use emotional information to predict consumers' emotions from their facial expressions (Hwang & Park, 2020). All these data collection and usage types can bring questions related to consumer privacy. What about the underlying reasons for consumer privacy concerns? According to researchers, privacy concerns come from multiple sources such as information collection, unauthorized use of information, and information access by third parties (Malhotra, Kim, & Agarwal, 2004; Smith, Milberg & Burke, 1996).

Data Collection, Storage, and Use: For instance, Google Home Mini has been reported to record intermittently during the day without the knowledge of its user and transfer the data back to Google (Forrest, 2017) that trigger consumer privacy concerns from the point of data collection and use. Similarly, firms' use of images and personal information obtained from a user voluntarily can even uncover copyright and ownership problems in the future (Lobera, Fernández Rodríguez & Torres-Albero, 2020). Moreover, firms' usage of consumer data for reasons different from consumers were informed while collecting the data is a way to violate consumer privacy (Davenport et al., 2020; Du & Xie, 2021; Tucker, 2018).

In addition to the problems related to the collection and use of data, we can mention issues regarding the storage of data. First of all, data can be stored longer than it was claimed to be stored, and second of all, some parts of the data can include information about other people (Davenport et al., 2020; Tucker, 2018). For instance, some neighbors might have concerns about Amazon's doorbells with cameras, the Ring device, recording their data without their permission.

Access of Data by Third Parties: Another concern related to privacy is the access of consumer data by third parties. Consumers question whether firms will sell the data collected through AI-enabled products to third parties without the users' permission (Davenport et al., 2020). Moreover, what will happen if the stored data are requested by the government or by the police in the future (Jones, 2018). Using both the data from the government and AI-enabled device companies, the higher level of surveillance in society is also another concern about consumer privacy (Lyon, 2014).

Cybersecurity is another significant concern about AI. Because of possible data breaches, malicious parties can get consumers' confidential information (Du & Xie, 2021). The results of data breaches can be devastating, such as the stolen passwords or the hacked bank accounts (Lobera, Fernández Rodríguez

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

& Torres-Albero, 2020). As the data collection of companies increase, especially with the help of AI-enabled products, the potential risk of a data breach also increases (Du & Xie, 2021). There can be two types of cyber-attacks related to AI; the first type is the attacks on AI and the second one is about the attacks using AI (Hwang & Park, 2020). The first type is the attacks on AI; they are about hacking an AI. Examples can be hacking a self-driving vehicle or hacking Cyber-Physical Systems such as smart governance, smart transportation (Hwang & Park, 2020). The other one is about the attacks using AI, such as speech synthesis, in which an AI is taught to mimic a person's voice. In this way, financial fraud can be committed through telephone calls (Kaloudi & Li, 2020).

So what can firms do to overcome privacy and cybersecurity concerns? Du and Xie (2021) offer three fundamental steps for firms to address consumer privacy concerns. The first one is to offer fair and transparent privacy policies and communicate their privacy policies to consumers with a clear and easy-to-understand communication. The second one is to explain what benefits consumers will get when sharing their data. Lastly, the third one is to give consumers more control over their data. Also, for cybersecurity concerns, firms should take precautions such as investing in the newest security technologies (Du & Xie, 2021) and communicate to their consumers that they are taking the security of their data very seriously. For the times data breaches occur, companies should have a plan to minimize the damage to their consumers (Du & Xie, 2021).

Algorithmic Biases

AI-enabled products learn based on machine learning. In this learning process, training data sets that are based on human data are used, and AI develops algorithms based on this training data. In other words, AI learns based on human data. Thus, while AI learns how to decide like a human, it also learns to make the exact same errors that humans do. Sometimes these errors lead AI to do human discrimination (Hwang & Park, 2020). For example, Amazon realized that its AI recruitment tool discriminates against female candidates and abandoned this project (Hamilton, 2018). Banks using AI are also accused of discrimination. A customer complains about being given a higher amount of a credit limit than his wife, although they have the exact same financial backgrounds (Puntoni et al., 2021). Amazon's Rekognition falsely matched innocent people with people who had committed a crime, and the number of people of color was higher in this false match (Snow, 2018).

When we think about the source of these algorithmic biases, we can understand the importance of the training data because AI biases come from the unbalanced and biased training data. The possible imbalances in the training data, such as gender, race, and ethnicity, cause biases in the AI-enabled products' algorithms (Du & Xie, 2021). For example, the reason behind the Amazon AI recruitment tool's discrimination against female candidates was the usage of a data set based on previous applicants to develop the algorithm, and the previous applicants being male (Davenport et al., 2020). This also shows us that, either intentionally or unintentionally, we transfer our biases to AI-enabled products (Letheren, Russell-Bennett & Whittaker, 2020). Just as the parents or teachers transmit their ideas to a child, we transmit our biases to AI (Letheren, Russell-Bennett & Whittaker, 2020; Sumioka, Nakae, Kanai, & Ishiguro, 2013).

So, what can be done to overcome artificial intelligence biases? First of all, companies should be careful about their training data sets; they should be sure that the training data sets are diverse enough to represent different human segments. Using balanced data sets from the point of gender, ethnicity, or race should be the first step to overcome algorithmic biases (Du & Xie, 2021). As the second step, companies

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

should continuously test their algorithms to learn about the possible algorithmic biases they might have (Puntoni et al., 2021). Looking at the artificial intelligence biases from the consumers' perspective shows us that the situation can be devastating, and consumers can feel misunderstood when they encounter an AI product making biased predictions about them. Thus, to overcome consumers' resistance to artificial intelligence and consumers' feelings of misunderstood due to miscategorization, companies can allow consumers to validate the categorizations and inferences of artificial intelligence (Puntoni et al., 2021).

Vulnerable Consumers

While AI spread into all our lives, there are many people whose access to digital devices or the knowledge and ability to use AI technology are limited. These people can be among poor, elderly, and unwell consumer groups and can be excluded and alienated in the marketplace (Hwang & Park, 2020). Also, the possibility that human services would be more exclusive and expensive in the future brings the concern that vulnerable groups such as consumers experiencing poverty might be pushed to receive service only from AI representatives (Letheren, Russell-Bennett & Whittaker, 2020). For consumers who need human-to-human interaction and who do not want to interact with an AI agent, alienation would be inevitable too (Letheren, Russell-Bennett & Whittaker, 2020). Thus, firms should be flexible enough to offer a 'human touch' for consumers who are either resistant to AI or experiencing problems with AI (Puntoni et al., 2021).

AI in Work Environment and Unemployment Concerns

Another category for consumers' concerns is the concerns related to unemployment, working conditions of AI, co-working of human-AI employees, and consumers' concerns about delegating their own tasks to AI. First of all, from the point of unemployment, we can see that both people and experts worry about the possible future job losses because of the AI job replacements. Today, AI is being used in a variety of jobs from low-skilled ones to high-skilled ones. For example, we see AI taking orders from customers in the form of chatbots or performing the work of consultants, lawyers, or doctors (Meltzer, 2014; Susskind & Susskind, 2016). AI is also expected to be more intelligent with more capabilities in the future which will lead to more job replacements (Du & Xie, 2021). The reports of respected institutions and consultancy firms such as the World Economic Forum, McKinsey Global, and PwC support that unemployment caused by AI can be a problem in the future that we have to deal with (Bughin et al., 2017; PWC, 2018; Upchurch, 2018). For example, McKinsey predicted that; by 2030, up to 375 million workers might have to switch occupations and learn new skills (Manyika et al., 2017). Research shows that consumers' concerns about losing their jobs are an important factor that leads them to resist AI (Lobera, Fernández Rodríguez & Torres-Albero, 2020).

Moreover, as AI disseminates in the workplace, concerns about the consequences and the possible problems of co-working of AI with human employees emerge (De Keyser et al., 2019; Robinson et al., 2020). In line with this concern, literature underlines the importance of designing AI with needed cognitive capabilities to work with human teammates (Dwivedi et al., 2021). Another concern focuses on the possible work conditions of AI. AI can be the dreamed workforce for businesses since it can work for long hours without rest, does not ask for higher wages, and can accomplish even the most boring or dangerous jobs (Letheren, Russell-Bennett & Whittaker, 2020). It is criticized that this type of usage of AI is a new form of slavery (Petersen, 2007; Richardson, 2016), and the introduction of AI rights can

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

be necessary (Letheren, Russell-Bennett, & Whittaker, 2020). The last issue that we find appropriate to discuss under this title is about consumers delegating their own tasks to AI. Sometimes consumers can be happy and feel empowered in delegating some of their tasks to AI. However, some of the other times delegating can make them feel replaced (Puntoni et al., 2021), can create negative identity perceptions (Leung, Paolacci & Puntoni, 2018), and negatively influence their self-worth and self-efficacy (Puntoni et al., 2021).

So, what can be done to overcome these concerns? Though it is mentioned that some jobs will be completely replaced by AI and disappear, experts believe that new jobs will appear (Manyika et al., 2017). Also, firms are not completely replacing human employees but supporting their capabilities with AI technology (Davenport et al., 2020). Thus, the communication of these ideas to people can help to reduce the concerns. Furthermore, firms should give opportunities for their employees for life-long learning and reskilling (Du & Xie, 2021). Finally, AI should be designed in ways to help consumers protect their identity perception while they are delegating their tasks (Puntoni et al., 2021).

AI and Ethical Decision Making

The ethical decision-making of AI creates another concern. People have to trust AI that it can follow moral rules in decision-making. The famous example here is the trolley problem, where AI systems at some point should have to make decisions like to either sacrifice one person for saving a larger number of people or the opposite (Thomson, 1976). In an attempt to answer these kinds of ethical dilemma and satisfy consumers' concerns about the morality of the decisions of AI, ethics for machines have to be created (Anderson & Anderson, 2011). Hence, organizations and researchers have directed their efforts towards examining and creating ethics for machines. For example, the European Commission has created a guideline named "Ethics Guidelines for Trustworthy AI" (European Commission, 2019), and a new research field named "machine ethics" or "machine morality" has emerged (Deng, 2015; Lin, Abney & Bekey 2012; Malle, 2016; Wallach, Franklin & Allen 2010).

Another question that comes to mind concerning AI ethics can be; what do ethical standards or machine morality mean? Ethical standards are the "guidelines based on morals and used to inform behaviour in situations that may not be covered by the law" (Letheren et al, 2020, p. 227) and machine morality deals with the ethical standards of artificial intelligence and, more specifically, autonomous machines (Malle, 2016). Since the laws about AI decisions and the consequences of AI decisions are not fully established yet, setting ethical standards regarding the decisions of AI is very important. These ethical standards can be significant guides both for marketers and AI developers (Letheren et al., 2020).

In setting ethical guidelines, consumers' perspectives and expectations related to AI morality come to the scene. It is important to understand consumers' perceptions and expectations about the moral norms of AI. What sort of moral norms are people expecting from these AI agents? Whether their moral norm expectations from AI are similar to their expectations from humans, or are they different? Understanding the answers to these questions is so crucial in designing future AI agents (Gill, 2020). Thus, in literature, we see much research focusing on these questions. Especially, the decisions of autonomous vehicles in life-and-death decision scenarios have taken a significant research interest (Awad et al. 2018; Bigman & Gray 2020; Bonnefon, Azim & Iyad 2016; Faulhaber et al. 2019; Gill, 2020). For example, Gill (2020) tried to reveal whether consumers expect an autonomous vehicle carrying them, either to swerve off the road and maybe kill them but save a pedestrian or just the opposite. Moreover, consumers' expectations about the moral norms of autonomous vehicles can differ based on the target characteristics (one vs. five

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

pedestrians, child vs. adult pedestrian), severity of the harm, or the perspective they are taking (perspective of the passenger in autonomous vehicle vs. pedestrian) (Gill, 2020). Meanwhile, manufacturers try to find their own solutions; while some of them announce that their autonomous vehicles will protect their passenger even at the expense of others (Sorrel, 2016), others try to position themselves into a safer area and underline consumers' customization option based on their own preferences (Contissa, Lagioia & Sartor, 2017). In sum, the research, debates, and consumers' concerns about AI ethical decision-making seem to continue for a longer period of time.

FUTURE RESEARCH DIRECTIONS

AI has been evolving since the 1940s, and today we can see it both in the applications of businesses and directly in consumers' lives. One of the ways companies use AI is as their employee that interacts directly with consumers in the form of chatbots, sales assistants, financial assistants, or robots. This type of AI usage is the one that directly touches the consumer, and we see these types of usages in a variety of sectors such as real estate, finance, health care, finance, beauty, fashion, and hospitality. Another way businesses use AI is as helping them in their tasks without a direct consumer touch, such as recommending systems that analyze consumers' previous actions and recommend them products or services or as the systems used in healthcare to evaluate radiological images or pathology slides. Additionally, businesses can offer AI to the consumers as the product or the service itself, such as voice-controlled digital assistants or companion robots.

From all the pervasive usages of AI by consumers and businesses, it is clear that a new marketplace has been started to be shaped. The human-to-human interaction was at the core of the traditional service encounter. Nevertheless, today, we have started to see the examples of AI agents representing the company in the form of chatbots or representing the consumer in the form of digital assistants, and this brings AI-to-human or AI-to-AI interactions to the marketing scene. Moreover, all these consumer interactions with AI bring new experiences for the consumers, which in turn lead them to feel positive such as served, understood, empowered, connected, or negative such as exploited, misunderstood, replaced, or alienated (Puntoni et al., 2021). In these interactions, not only the consumers are having an experience, but also the AI is having an object experience (Hoffman & Novak, 2018).

Thus, future research should try to find answer to some questions that have not examined previously. For example, how will firms treat their new AI customers, how will these new AI customers affect firms' marketing strategy, will firms target human customers or target the AI customer in their marketing strategy? For instance, whether a firm trying to promote its new brand will target the digital assistant of the consumer or the consumer himself/herself? The interactions and experiences within AI customers are another area that needs attention such that; what will happen to the dynamics of word of mouth when AI customers write the reviews about the products, services or brands they used and liked or do not liked. Will customers still believe these reviews or not? Moreover, the interactions and experiences within AI employees or the interactions and experiences between AI and human employees in marketing departments of firms should be analyzed too. As Hoffman and Novak (2018) underlined, the object experience of the AI should also be considered. Is it really possible to think about the experience of an AI agent and whether this object experience will affect an AI customer to shop from an online store or not based on its previous experience? In other words, whether the object experience of AI customers

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

will affect their repurchase intentions and brand loyalty? What would marketers do to positively affect the experience of the AI customers?

CONCLUSION

As a result of consumers' interactions and experiences with AI, it is expected that consumers might build relationships with the AI. So, it is crucial to understand how consumers perceive their relationships with the AI and what roles they assign to themselves or the AI in the relationship. But, our literature review has shown us that research focusing on these questions are scarce (Schweitzer et al., 2019; Huang et al., 2021). Although Schweitzer et al. (2019) showed us that consumers describe their relationship with their digital assistant as; AI being the servant, master, or a partner, more research is needed in this fruitful area.

The usage of AI has been spread into the life of consumers directly or indirectly and has many benefits, such as doing many tasks on behalf of consumers. Nevertheless, people can still resist getting services from AI and prefer human service instead of AI service. There is a vast amount of research examining consumers' attitudes toward and resistance to AI. It has been revealed in the literature that the individual factors such as the scientific knowledge of people, their openness to innovation, their risk perceptions have an effect on their acceptance of AI (Castelo & Ward 2016; Lobera, Fernández Rodríguez & Torres-Albero, 2020). Also, consumers' identity issues, their uniqueness preference in symbolic consumption contexts, and the characteristics and consequences of the tasks can affect their preference to choose to get service from an AI agent. What other factors lead consumers to prefer AI service or resist AI service is another research area that needs more examination.

Finally, the chapter examined the major concerns both consumers, society and researchers have about AI. These were about privacy and cybersecurity, AI biases, discrimination of vulnerable consumers, possible unemployment issues because of AI, and the ethicality of AI decisions. New research is needed in understanding the consumers' and society's concerns about AI, and these concerns should be taken into consideration in the design of AI systems, and marketers should follow these concerns in using and promoting their AI-enabled systems and products.

REFERENCES

Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *The Journal of Consumer Research*, *34*(4), 468–479. doi:10.1086/518544

Akhter, S. H. (2014). Privacy concern and online transactions: The impact of internet self-efficacy and internet involvement. *Journal of Consumer Marketing*, *31*(2), 118–125. doi:10.1108/JCM-06-2013-0606

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2020). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, *114*, 106548. doi:10.1016/j.chb.2020.106548 PMID:32905175

Amelia. (n.d.). *Amelia: The most human AI for the Enterprise*. https://amelia.com/amelia/

Anderson, M., & Anderson, S. L. (Eds.). (2011). *Machine ethics*. Cambridge University Press. doi:10.1017/CBO9780511978036

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Antonio, V. (2018, July 30). How AI is changing sales. *Harvard Business Review*. https://hbr.org/2018/07/how-ai-is-changing-sales

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64. doi:10.103841586-018-0637-6 PMID:30356211

Bhatia, R. (2016, December 12). Say hello to Simi – a chatbot that makes your cocktail. *Analytics India Magazine*. https://analyticsindiamag.com/say-hello-simi-chatbot-makes-cocktail/

Bigman, Y. E., & Gray, K. (2020). Life and Death Decisions of Autonomous Vehicles. *Nature*, *579*(7797), E1–E2. doi:10.103841586-020-1987-4 PMID:32132695

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The Social Dilemma of Autonomous Vehicles. *Science*, *352*(6293), 1573–1576. doi:10.1126cience.aaf2654 PMID:27339987

Bughin, J., Hazan, E., Ramaswamy, S., Allas, T., Dahlström, P., Henke, N., & Trench, M. (2017). Artificial intelligence: the next digital frontier? *McKinsey Global Institute*. https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. doi:10.1037/0033-2909.125.3.367

Castelo, N. (2019). *Blurring the line between human and machine: Marketing artificial intelligence* (Doctoral dissertation). Columbia University Academic Commons. doi:10.7916/d8-k7vk-0s40

Castelo, N., Bos, M., & Lehman, D. (2018). *Consumer adoption of algorithms that blur the line between human and machine*. Graduate School of Business: Columbia University Working Paper.

Castelo, N., & Ward, A. (2016). Political affiliation moderates attitudes towards artificial intelligence. In P. Moreau & S. Puntoni (Eds.), *NA - Advances in Consumer Research* (pp. 723–723). Association for Consumer Research.

Contissa, G., Lagioia, F., & Sartor, G. (2017). The ethical knob: Ethically-customizable automated vehicles and the law. *Artificial Intelligence and Law*, *25*(3), 365–378. doi:10.100710506-017-9211-z

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. doi:10.100711747-019-00696-0

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, *96*(1), 108–116.

De Keyser, A., & Köcher, S., Alkire (née Nasr), L., Verbeeck, C., & Kandampully, J. (2019). Frontline service technology infusion: Conceptual archetypes and future research directions. *Journal of Service Management*, *30*(1), 156–183. doi:10.1108/JOSM-03-2018-0082

Deng, B. (2015). Machine ethics: The robot's dilemma. *Nature*, *523*(7558), 24–26. doi:10.1038/523024a PMID:26135432

DesJardins, J. (2014). An introduction to business ethics. New York, NY: McGraw-Hill/Irwin.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology. General*, *144*(1), 114–126. doi:10.1037/xge0000033 PMID:25401381

Du, S., & Xie, C. (2021). Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, *129*, 961–974. doi:10.1016/j.jbusres.2020.08.024

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. doi:10.1016/j.ijinfomgt.2019.08.002

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. doi:10.1038/nature21056 PMID:28117445

European Commission. (2019). *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/futurium/en/ai-alliance-consultation

Evans, M. (2019). Build a 5-star customer experience with artificial intelligence. *Forbes*. https://www.forbes.com/sites/allbusiness/2019/02/17/customer-experience-artificial- intelligence/#1a30ebd415bd

Faulhaber, A. K., Dittmer, A., Blind, F., Wächter, M. A., Timm, S., Sütfeld, L. R., Stephan, A., Pipa, G., & König, P. (2019). Human decisions in moral dilemmas are largely described by utilitarianism: Virtual car driving study provides guidelines for autonomous driving vehicles. *Science and Engineering Ethics*, *25*(2), 399–418. doi:10.100711948-018-0020-x PMID:29357047

Feng, C. M., Park, A., Pitt, L., Kietzmann, J., & Northey, G. (2020). Artificial intelligence in marketing: A bibliographic perspective. *Australasian Marketing Journal*. doi:10.1016/j.ausmj.2020.07.006

Forrest, C. (2017, October 11). Google Home Mini spied on user 'thousands of times a day,' sent recordings to Google. *TechRepublic*. https://www.techrepublic.com/ article/google-home-mini-spied-on-user-thousands-of- times-a-day-sent-recordings-to-google/

Gallagher, J. (2017, January 26). Artificial intelligence 'as good as cancer doctors'. *BBC News*. https://www.bbc. com/news/health-38717928

Gill, T. (2020). Blame it on the self-driving car: How autonomous vehicles can alter consumer morality. *The Journal of Consumer Research*, *47*(2), 272–291. doi:10.1093/jcr/ucaa018

Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, *115*, 106607. doi:10.1016/j.chb.2020.106607

Granulo, A., Fuchs, C., & Puntoni, S. (2020). Preference for human (vs. robotic) labor is stronger in symbolic consumption contexts. *Journal of Consumer Psychology*, *31*(1), 72–80. doi:10.1002/jcpy.1181

Greenwald, T. (2011). How smart machines like iPhone 4S are quietly changing your industry. *Forbes*. https://www.forbes.com/sites/tedgreenwald/2011/10/13/how-smart-machines-like-iphone-4s-are-quietly-changing-your-industry/#46547361598f

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gustafsod, P. E. (1998). Gender Differences in risk perception: Theoretical and methodological perspectives. *Risk Analysis*, *18*(6), 805–811. doi:10.1111/j.1539-6924.1998.tb01123.x PMID:9972583

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*(4), 5–14. doi:10.1177/0008125619864925

Hamilton, I. A. (2018, October 13). Why it's totally unsurprising that Amazon's recruitment AI was biased against women. *Insider*. https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10

Henn na Hotel. (n.d.). https://www.h-n-h.jp/en/concept

Hinchliffe, E. (2017, February 22), IBM's Watson will help diagnose heart disease when doctors may have missed the signs. *Mashable*. https://mashable.com/2017/02/22/ibm-watson-clinical-imaging-cardiology/

Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the internet of things: An assemblage theory approach. *The Journal of Consumer Research*, *44*(6), 1178–1204. doi:10.1093/jcr/ucx105

Hong, J. W. (2020). Why Is Artificial Intelligence Blamed More? Analysis of Faulting Artificial Intelligence for Self-Driving Car Accidents in Experimental Settings. *International Journal of Human-Computer Interaction*, *36*(18), 1768–1774. doi:10.1080/10447318.2020.1785693

Huang, B., Laporte, S., Sénécal, S., & Sobol, K. (2021), Partner or servant? The effect of relationship role on consumer interaction with artificial intelligence. *AMA Winter Academic Conference Proceedings*, *32*, 382-384.

Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. doi:10.1177/1094670517752459

Hwang, H., & Park, M. H. (2020). The Threat of AI and Our Response: The AI Charter of Ethics in South Korea. *Asian Journal of Innovation & Policy*, *9*(1), 56–78.

Jarek, K., & Mazurek, G. (2019). Marketing and Artificial Intelligence. *Central European Business Review*, *8*(2), 46–55. doi:10.18267/j.cebr.213

Jones, V. K. (2018). Voice-activated change: Marketing in the age of artificial intelligence and virtual assistants. *Journal of Brand Strategy*, *7*(3), 233–245.

Kaloudi, N., & Li, J. (2020). The AI-based cyber threat landscape: A survey. *ACM Computing Surveys*, *53*(1), 1–34. doi:10.1145/3372823

KLM. (2016, March 29). *Spencer robot completed tests guiding KLM passengers at Schiphol*. https://news.klm.com/spencer-robot-completed-tests-guiding-klm-passengers-at-schiphol/

Klmcare. (n.d.). *Meet KLM Care-E*. http://klmcare-e-entry.com

Lanctot, R. (2017). *Accelerating the Future: The Economic Impact of the Emerging Passenger Economy.* Strategy Analytics Report. https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/05/passenger-economy.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Leachman, S. A., & Merlino, G. (2017). Medicine: The Final Frontier in Cancer Diagnosis. *Nature*, *542*(7639), 36–38. doi:10.1038/nature21492 PMID:28150762

Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, *80*(6), 69–96. doi:10.1509/jm.15.0420

Letheren, K., Russell-Bennett, R., & Whittaker, L. (2020). Black, white or grey magic? Our future with artificial intelligence. *Journal of Marketing Management*, *36*(3-4), 216–232. doi:10.1080/0267257X.2019.1706306

Leung, E., Paolacci, G., & Puntoni, S. (2018). Man versus machine: Resisting automation in identity-based consumer behavior. *JMR, Journal of Marketing Research*, *55*(6), 818–831. doi:10.1177/0022243718818423

Lin, P., Abney, K., & Bekey, G. A. (Eds.). (2012). *Robot ethics: The ethical and social implications of robotics*. MIT Press.

Lobera, J., Fernández Rodríguez, C. J., & Torres-Albero, C. (2020). Privacy, values and machines: Predicting opposition to artificial intelligence. *Communication Studies*, *71*(3), 448–465. doi:10.1080/10510974.2020.1736114

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *The Journal of Consumer Research*, *46*(4), 629–650. doi:10.1093/jcr/ucz013

Longoni, C., & Cian, L. (2020). Artificial intelligence in utilitarian vs. hedonic contexts: The "word-of-machine" effect. *Journal of Marketing*. doi:10.1177/0022242920957347

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, *38*(6), 937–947. doi:10.1287/mksc.2019.1192

Lyon, D. (2014). Surveillance, snowden, and big data: Capacities, consequences, critique. *Big Data & Society*, *1*(2), 1–13. doi:10.1177/2053951714541861

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing–Connecting computing power to human insights. *International Journal of Research in Marketing*, *37*(3), 481–504. doi:10.1016/j.ijresmar.2020.04.005

MacInnis, D. J., & Folkes, V. S. (2010). The disciplinary status of consumer behavior: A sociology of science perspective on key controversies. *The Journal of Consumer Research*, *36*(6), 899–914. doi:10.1086/644610

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns: The construct, the scale, and a causal model. *Information Systems Research*, *15*(4), 336–355. doi:10.1287/isre.1040.0032

Malle, B. F. (2016). Integrating Robot Ethics and Machine Morality: The Study and Design of Moral Competence in Robots. *Ethics and Information Technology*, *18*(4), 243–256. doi:10.100710676-015-9367-8

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015), Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *HRI'15: Proceedings of the Tenth Annual ACM/ IEEE International Conference on Human-Robot Interaction*. New York, NY: ACM. 10.1145/2696454.2696458

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., & Sanghvi, S. (2017). *What the future of work will mean for jobs, skills, and wages.* McKinsey Global Institute Report. https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955), A proposal for the Dartmouth summer research project on artificial intelligence, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Meltzer, T. (2014, June 15). Robot doctors, online lawyers and automated architects: The future of the professions. *The Guardian*. https://www.theguardian.com/technology/2014/jun/15/robot-doctors-

Mintz, Y., & Brodie, R. (2019). Introduction to artificial intelligence in medicine. *Minimally Invasive Therapy & Allied Technologies*, *28*(2), 73–81. doi:10.1080/13645706.2019.1575882 PMID:30810430

Morgan, B. (2018). How Amazon has reorganized around artificial intelligence and machine learning. *Forbes*. https://www.forbes.com/sites/blakemorgan/2018/07/16/how-amazon-has-re-organized-around-artificial-intelligence-and-machine-learning/?sh=40ca07877361

Müller, V. C., & Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In Fundamental Issues of Artificial Intelligence (555-572). Springer. doi:10.1007/978-3-319-26485-1_33

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance through Text Mining. *Marketing Science*, *31*(3), 521–543. doi:10.1287/mksc.1120.0713

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. *California Management Review*, *61*(4), 156–185. doi:10.1177/0008125619859318

Pantano, E., & Pizzi, G. (2020). Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis. *Journal of Retailing and Consumer Services*, *55*(102096), 1–9. doi:10.1016/j.jretconser.2020.102096

Paschen, J., Kietzmann, J., & Kietzmann, T. C. (2019). Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. *Journal of Business and Industrial Marketing*, *34*(7), 1410–1479. doi:10.1108/JBIM-10-2018-0295

Petersen, S. (2007). The ethics of robot servitude. *Journal of Experimental & Theoretical Artificial Intelligence*, *19*(1), 43–54. doi:10.1080/09528130601116139

Poole, D. L., & Mackworth, A. K. (2010). *Artificial intelligence: Foundations of computational agents*. Cambridge University Press. doi:10.1017/CBO9780511794797

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, *85*(1), 131–151. doi:10.1177/0022242920953847

PWC. (2018). *Will robots really steal our jobs? An international analysis of the potential long term impact of automation*. PWC. https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/impact_of_automation_on_jobs.pdf

Revieve. (n.d.). *AI Skincare Advisor*. https://www.revieve.com/solutions/skincareadvisor

Richardson, K. (2016). Sex robot matters: Slavery, the prostituted, and the rights of machines. *IEEE Technology and Society Magazine*, *35*(2), 46–53. doi:10.1109/MTS.2016.2554421

Robinson, S., Orsingher, C., Alkire, L., De Keyser, A., Giebelhausen, M., Papamichail, K. N., Shams, P., & Temerak, M. S. (2020). Frontline encounters of the AI kind: An evolved service encounter framework. *Journal of Business Research*, *116*, 366–376. doi:10.1016/j.jbusres.2019.08.038

Romano, A. (2020, July 5). This hotel has a robot named Rosé that will deliver wine to your room Without human contact. *Travel and Leisure*. https://www.travelandleisure.com/food-drink/wine/rose-robot-room-service-hotel-trio-california

Rossi, F. (2018). Building trust in artificial intelligence. *Journal of International Affairs*, *72*(1), 127–134.

Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. Pearson Education Limited.

Russo, S. J. (2020). Is de-identification of personal health information in the age of artificial intelligence a reality or a noble myth? *Journal of Health Care Compliance*, (March–April), 55–59.

Rust, R. T. (2020). The future of marketing. *International Journal of Research in Marketing*, *37*(1), 15–26. doi:10.1016/j.ijresmar.2019.08.002

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management*, *35*(7-8), 693–715. doi:10.1080/0267257X.2019.1596970

Silverstein, S. (2018). REX real estate: Home showings-AI / Voice / Chatbot. *Behance*. https://www.behance.net/gallery/64004179/REX-Real-Estate-Home-Showings-AI-Voice-Chatbot

Smith, J. H., Milberg, S. J., & Burke, J. B. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *Management Information Systems Quarterly*, *20*(2), 167–196. doi:10.2307/249477

Snow, J. (2018, July 26). Amazon's face recognition falsely matched 28 members of congress with mugshots. *ACLU*. https:// www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Sorrel, C. (2016, November 13). Self-driving Mercedes will be programmed to sacrifice pedestrians to save the driver. *Fast Company*. https://www.fastcompany.com/3064539/self-driving-mercedes-will-be-programmed-to-sacrifice-pedestrians-to-save-the-driver

Sumioka, H., Nakae, A., Kanai, R., & Ishiguro, H. (2013). Huggable communication medium decreases cortisol levels. *Scientific Reports*, *3*(1), 1–6. doi:10.1038rep03034 PMID:24150186

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Susskind, R., & Susskind, D. (2016, October 11). Technology will replace many doctors, lawyers, and other professionals. *Harvard Business Review*. https://hbr.org/2016/10/robots-willreplace-doctors-lawyers-and-other-professionals

Thales. (2018, January 8). KFC Use Facial Recognition for Payment in China. *Thales*. https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/magazine/kfc-use-facial-recognition-payment-china

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, *59*(2), 204–217. doi:10.5840/monist197659224 PMID:11662247

Tirunillai, S., & Tellis, G. J. (2014). Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *JMR, Journal of Marketing Research*, *51*(4), 463–479. doi:10.1509/jmr.12.0106

Tucker, C. (2018). Privacy, algorithms, and artificial intelligence. In *The Economics of Artificial Intelligence: An Agenda* (pp. 423–438). University of Chicago Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *49*(236), 433–460. doi:10.1093/mind/LIX.236.433

Tussyadiah, I. (2020). A review of research into automation in tourism: Launching the Annals of Tourism Research Curated Collection on Artificial Intelligence and Robotics in Tourism. *Annals of Tourism Research*, *81*, 102883. doi:10.1016/j.annals.2020.102883

Upchurch, M. (2018). Robots and ai at work: The prospects for singularity. *New Technology, Work and Employment*, *33*(3), 205–218. doi:10.1111/ntwe.12124

Upstart. (n.d.). https://www.upstart.com/about#

Vincent, J. (2017, January 4). The UK's national health service is testing out a medical chatbot as a non-emergency helpline. *The Verge*. https://www.theverge.com/2017/1/4/14168312/uk-nhs-babylon-medical-chatbot-helpline

Voice2Biz. (n.d.). *Google Home and Alexa Talk to aach other & have a conversation*. https://www.voice2biz.com/google-home-and-alexa-talk-to-each-other-have-a-conversation/

Wallach, W., Franklin, S., & Allen, C. (2010). A Conceptual and Computational Model of Moral Decision Making in Human and Artificial Agents. *Trends in Cognitive Sciences*, *2*(3), 454–485. PMID:25163872

Wiggers, K. (2019, March 19). Ojo Labs raises $45 million to develop a chatbot for real estate. *VentureBeat*. https://venturebeat.com/2019/03/19/ojo-labs-raises-45-million-to-develop-a-chatbot-for-real-estate/

Wirth, N. (2018). Hello marketing, what can artificial intelligence help you with? *International Journal of Market Research*, *60*(5), 435–438. doi:10.1177/1470785318776841

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

AI Biases: Bias in the AI algorithm, which can lead to prejudiced AI decisions such as discrimination based on gender or ethnicity.

AI Ethical Decision Making: AI that follows moral rules in decision making.

Artificial Intelligence (AI): Computer systems that mimic the intelligence of humans. These computer systems perceive the world around them and take action accordingly.

Consumer Algorithm Aversion: Consumers' preference of humans over algorithms in a variety of tasks.

Hybrid AI: AI systems that integrate multiple weak AI systems such as Google's DeepMind AlphaGo or IBM's Watson.

Strong AI: AI type that is more powerful than weak AI. These systems have context-awareness, think holistically, give context-specific responses, and able to complete more complex tasks.

Weak AI: AI type that analyzes large amounts of data and can be trained to solve specific problems. They are powerful in their specifically trained area but not as flexible as human intelligence in other areas.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 17
When Luxury Vinous-Concept Hotel Meets Premium Wine Brands:
An Exploratory Study on Co-Branding

Lisa Ferraz
University of Aveiro, Portugal

Helena Nobre
https://orcid.org/0000-0002-7724-5204
University of Aveiro, Portugal

Belém Barbosa
https://orcid.org/0000-0002-4057-360X
University of Porto, Portugal

ABSTRACT

Co-branding in the hospitality luxury sector is still understudied in the literature. This study aims at tackling this gap through the analysis of a case of a co-branding strategy between a vinous-concept luxury hotel in Portugal and premium wine brands of domestic producers. Fourteen in-depth interviews with managers of the luxury hotel and wine brand partners supported the exploratory research. This chapter represents a case of qualitative data application to an underestimated topic in the literature from the managers' point of view. The study offers evidence on the benefits for both parties, reasons for adopting co-branding, and partners' selection attributes. The improvement of brand image emerges as one of the main advantages of co-branding with a luxury hotel. Based on the literature review and the interviews with managers, the study proposes a set of hypotheses to be tested in future research. This chapter provides interesting cues for academics and practitioners.

DOI: 10.4018/978-1-7998-6985-6.ch017

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Co-branding practices have been increasingly adopted by practitioners in several fields, including advertising, luxury market, products, and retail stores (e.g., Abratt & Motlana, 2002; Dahlstrom & Dato-on, 2004; Moon & Sprott, 2016), and also in the hospitality industry, namely restaurant chains (Boone, 1997), chefs, retail, local wine, designers, and other celebrity brands (Guillet & Tasci, 2012). Co-branding strategy gained popularity in the academic literature due to its ability to reinforce brand image (Lee, 2014; Park, Jaworski, & MacInnis, 1986) and improve brand perceptions and positioning (Singh, P. Kalafatis, & Ledden, 2014). Thus, it is essential to identify strong and reputable partner brands, based on criteria such as image, quality, brand equity and loyalty (Guillet & Tasci, 2012). In the retail sector, co-branding partnerships enable the access to wider and different marketing segments, to increase sales, and foster relationships with customers (Wang, Soesilo, & Zhang, 2015). Co-branding implies sharing brand value, business competences and risks (Erevelles, Stevenson, Srinivasan, & Fukawa, 2008; Helmig, Huber, & Leeflang, 2008; Keller & Richey, 2006), and enables the maximization of distribution channels and customer relationship programs (Motion, Leitch, & Brodie, 2003; Wang et al., 2015).

At the same time, luxury brands are getting increasing attention from academics and practitioners, as the luxury market shows a clear resilience regarding economic crisis, being one that consistently grew in the last years (Nobre & Simões, 2019). It is expected that it keeps growing in the next decades, namely because of the Millennial generation who will reach their thirties and forties, becoming one of the most important drivers of the global economy. Moreover, globalization and economic downturns increased competition pressure and made harder for companies to ensure that luxury brands' perceived value correspond to their price level (Díaz-Bustamante, Carcelén, & Puelles, 2016; Tynan, McKechnie, & Chhuon, 2010). Strategies that help improving luxury brands' value and positioning are particularly relevant for business success and marketing effectiveness. Nueno and Quelch (1998) define luxury brands as "those whose ratio of functional utility to price is low while the ratio of intangible and situational utility to price is high" (p. 62). Besides expensiveness associations, luxury brands share some characteristics such as consistent quality and product excellence, exclusivity and uniqueness, tradition and heritage, hand-crafted, rare, and elitist (Kapferer, 1998; Vigneron & Johnson, 1999). Moreover, the concept of luxury is also understood in a socio-psychological context, embedded in culture and lifestyle (Okonkwo, 2009). Thus, luxury items and brands transmit strong and solid identities, associated with symbolic and emotional consumption.

We argue that one particularly interesting setting for co-branding is the luxury hotel brand. Both the attractiveness and the complexity of the luxury markets, especially in hospitality services, claim for sophisticated business strategies like co-branding. Surprisingly, extant literature has so far missed the focus on co-branding strategies in the luxury market. Although some studies in the literature present evidence on co-branding being used for hotel differentiation and positioning (Guillet & Tasci, 2012), most of them disregard co-branding under the light of luxury market and brands. This article aims to fill this gap by exploring how a co-branding strategy between a luxury hotel brand and product brands benefits each other, especially in the case of a product particularly relevant for the hotel brand concept.

The study analyses the case of a local luxury brand Hotel in Portugal, with a vinous-concept positioning, and its partnership with premium wine brands of domestic producers. Based on the perceptions of the hotel and wine brands' managers, the study aimed to identify the determinants for partner choice, and the impact of the co-branding strategy in brand image of both the hotel and the associated wine brands.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Qualitative data is particularly suitable to investigate new or quite unexplored phenomenon in the literature (Glaser & Strauss, 2006). Particularly, in-depth interview represents a common technique to collect data on consumers' perceptions and opinions (Goulding, 2005). This article represents a case of qualitative data application to an underestimated topic in the literature, but in contrast with most of previous research, the study advances literature on co-branding in luxury hospitality management from the managers' point of view.

CO-BRANDING STRATEGY

In the past decades, the brand partnerships of co-branding strategies became more and more popular (Riley, Charlton, & Wason, 2015; Wason & Charlton, 2015). Chiambaretto and Gurău (2017) define co-branding as a strategy that combines two or more independent brands in one offer. Brennan, Canning, and McDowell (2014) consider that a co-branding results from a cooperation between two or more companies comprising some defining aspect of the brand. Indeed, co-branding is often used interchangeably with other concepts such as brand alliances, comarketing, joint branding, branding and symbolic branding.

The main objectives of co-branding is to create mutual benefits for the involved brands (Egan, 2011), namely to improve their brand equity (Abratt & Motlana, 2002; Genc, 2010; Guillet & Tasci, 2012). This cooperation often involves several marketing domains such as advertising, product development and positioning, and distribution channel strategies (Grossman, 1997). Partner companies are expected to develop their individual competences and share resources (Shen, Choi, & Chow, 2017) to mutually reinforce their brand equities. Abratt and Motlana (2002) admit that co-branding main advantage is convincingly position the brands, which generates more sales in both existent and new market segments and increased brand value. Co-branding strategies are also expected to reduce the costs of introducing a new product (Helmig et al., 2008), and to provide competitive advantage and increased attractiveness namely due to the spillover effects (Erevelles et al., 2008).

Despite its advantages, co-branding can also generate negative impacts. Helmig et al. (2008) stress the need of adjustments from each partner, stating that success will depend on the commitment of all partners involved. Co-branding is, indeed, a complex (Helmig et al., 2008) and inflexible (Lee, 2014) strategy.

Brand Image, Brand Identity, and Co-Branding

Brand identity expresses the tangible and intangible characteristics of the brand that makes it unique (Kapferer, 2012). Aaker (1996) distinguishes brand image from brand identity by explaining that image refers to the subjective perspectives at a determined moment about the brand, while identity is how managers wish target audiences understand the brand, stressing that effectively expressing brand identity is a key factor for brand success. Motion et al. (2003) found that managers tend to establish relationships in order to redefine brand identity. Alliances with other brands contribute to the development of new identities (Balmer & Greyser, 2002). As brand partnerships may have a great impact on companies, they should be carefully evaluated (Öberg (2016).

According to Riley et al. (2015), consumers' perceptions on a brand alliance depend on the adequacy of the alliance, that is, the consistency of the brands in the partnership. Moreover, the less known brand tends to converge to the brand with more awareness (Abratt & Motlana, 2002), resulting in the transference of positive associations of the stronger brand to the other. This evidence is consistent with results

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

obtained by Geylani, Inman, and Hofstede (2008) who found that the partner with best performance has a positive impact on co-branding product perceptions. Still, Geylani et al. (2008) recommend choosing partners with similar performance, in order to avoid the perception of discrepancies by consumers.

Building a positive brand image can represent the main purpose of a co-brand strategy by reinforcing the image of the partner brands (Geylani et al., 2008), allowing the brand to set a different position toward the competition. In the hospitality sector, namely luxury hotels, value maximization, needs' fulfilment, and guests satisfaction are essential to develop a strong brand image (e.g., Dev, Morgan, & Shoemaker, 1995; Li, Yen, & Uysal, 2014; Su & Reynolds, 2017). Clearly, co-branding strategies require a rigorous evaluation of the partner involved (e.g., Guillet & Tasci, 2012; Helmig et al., 2008), as shown by the wine business sector, where the establishment of associations increases the possibility of more effective promotion and the creation of strong points-of-differentiation (Vrontis, Thrassou, & Czinkota, 2011).

The Importance of Co-Branding to Reinforce Brand Positioning

Brand positioning consists of the mission and the brand value proposition (Nobre, 2019) - functional, emotional, and self-expression benefits - that is actively communicated to the market (Aaker, 1996). It assumes importance in the formation of consumers' perceptions regarding subjective and objective brand attributes (Kapferer, 2012; Lewis, 1981) and provides companies with important competitive advantages (Malik, Naeem, & Munawar, 2012), by establishing differentiation and assuring consumer preference. Singh et al. (2014) found that consumers' perceptions on brand positioning are transferable to the other partner brands. Hence, co-branding strategies act as endorsers of other brands' quality. Wason and Charlton (2015) suggest that brand positioning is a fundamental criterion for partner selection, referring that brands with hedonic attributes are more able of achieving partnership benefits than brands more associated with functional attributes. Moreover, brand equity should also be carefully considered by functional brands, while hedonic brands need to further develop partner brands' associations (Wason & Charlton, 2015).

The hotel industry, in the main cities in Europe, is generally higly fragmented and under fierce competition of direct local hotel brands, global hotel chains and a myriad of diverse substitutes that pressure prices and impact market dynamics. This impels hotel brands, in particular upper-class hospitality services, for fighting for a unique position and a sustainable competitive advantage in the market (Li, Yen & Uysal, 2014). A co-branding strategy can be particularly suitable to build or reinforce a unique positioning for a hotel brand. A co-branding strategy might offer an opportunity of differentiation and of building a strong positioning, besides the risks and challenges that a partnership with other brands and their associated brand images can represent at the level of consumer acceptation and marketing communication consistency (Hadjicharalambous, 2006; cit. in Guillet & Tasci, 2012).

METHOD

Based on the contributions from extant literature, three research questions guided the present study:

RQ1: How do luxury hotel and local wine brands benefit from co-branding partnerships?

RQ2: How does a co-branding strategy impact on partners' brand image management?

RQ3: How do partnerships with prestige products influence luxury service brand positioning?

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Participants in the study: Luxury hotel brand managers

Participants	Gender	Age group	Education	Years in the Company	Job
Interviewee 1	F	36-45	Post-graduate	8	Wine Director
Interviewee 2	F	26-35	Post-graduate	8	Head Sommelier
Interviewee 3	F	18-25	Graduate	3	Marketing Assistant
Interviewee 4	F	26-35	Post-graduate	2	Brand Assistant
Interviewee 5	M	26-35	Post-graduate	2	Designer
Interviewee 6	M	46-55	Graduate	7	CMO

In order to tackle these research questions, we conducted a qualitative exploratory study, as co-branding in luxury hospitality management still lacks for strong validation from the literature. Qualitative data is "the best and richest for theorizing about social structures and social systems" (Glaser & Strauss, 2006, p.17). This study, in particular, focus on a co-branding partnership between wine brands and a vinous-concept luxury hotel brand, from now on referred to as the luxury hotel. Thus, the managers' opinions and convictions are crucial for understanding the partnerships between the brands they represent. All these aspects sustain the relevance of considering manager's narratives to investigate this case.

Located in Portugal, this hotel currently maintains a partnership with around 102 brands of wine producers, which helps maintaining the vinous-concept established by the luxury hotel in all its valences. Each of the hotel's rooms has a name of a wine brand and decoration according to its characteristics. Besides allowing the luxury hotel to be associated with any wine brand, the partnership includes having the partner wine brands sold at the hotel's store and website, and featured in their restaurant, including in exclusive wine dinner events.

We conducted 14 semi-structured interviews with managers of the luxury hotel (Table 1) and he wine brand partners (Table 2). Interviewees has access to full information on the study and gave their

Table 2. Participants in the study: Luxury hotel brand managers

Participants	Company size	Gender	Age group	Education	Years in the Company	Job
Company A	Small	M	36-45	Graduate	7	CEO
Company B	Small	F	46-55	Graduate	7	Sales manager
Company C	Large	F	26-35	Post-graduate	3	Marketing and Sales Manager
Company D	Medium	F	26-35	Post-graduate	3	Key Account Manager
Company E	Small	M	56-65	Graduated	12	CEO
Company F	Medium	F	46-55	Post-graduate	8	Marketing and Sales Manager
Company G	Medium	M	46-55	Post-graduate	3	CEO
Company H	Large	M	36-45	Undergraduate	20	Sales Manager

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

informed consent prior to participating in the study. We also followed other ethical principles for social sciences research, including ensuring that participation was voluntary, confidential and anonymous.

Invitations to participate in the study were initially sent by e-mail, along the interview guidelines and purpose of the study. We conducted six face-to-face interviews and, due to scheduling difficulties, one interview by telephone call, and the other seven by e-mail. The telephone and face-to-face interviews were recorded to allow transcription, with the participants' consent. Besides the lack of intimacy, connection and spontaneity that email interviews could represent, this question seems not to be so problematic when participants are managers (from the supply side) and can otherwise represent a strength of the study, as they encourage for more correctness, detail and rigour in the information provided. Interviews' duration ranged from 15 to 25 minutes. A set of categories gathered from the literature review guided content analysis, comprising co-branding strategy implementation (partner selection criteria and shared benefits), impact on partners' brands (positioning and image), and importance of co-branding partnerships as distribution channels.

RESULTS

Co-Branding Mutual Benefits

Co-branding mutual benefits were acknowledging by all participants in this study. Despite the strong positioning and image of the luxury hotel, and the opportunity it offers to their wine partners to reach the luxury customers, the Hotel's representatives recognize the mutual advantage provided to the involved brands, but with different intensity along the process. Indeed, gains for the Hotel were particularly visible during its launch stage, as in order to establish its vinous-concept (the basis of it differentiation) the Hotel benefited of the association with well-known high quality wine brands. For instance, the Hotel's Sommellier (Interviewee 2) noted that this was determinant for the vinous-concept Hotel's brand success:

The [Hotel] enjoys this advantage of having quality producers associated with the brand, that is, they also leverage us in the sense that we want to be an embassy of Portuguese wines (…) and the growth of our brand was due to having producers of recognized quality that joined us, gave credibility to the Hotel, since they were in the market for longer than us.

These benefits continue to flow, now that the Hotel's image is established. The Hotel's CMO (Interviewee 6) explains that "aligning with the best wine producers in Portugal helps reinforce the Hotel's concept, positioning the Hotel as an authority in the wine context".

Clearly, the success of this co-branding strategy relates to the ability of consistently providing benefits to partners and the congruity of the brands involved. As one of the large company partners referred, this is due to the shared values of the partners "both brands represent a common ideology related to luxury, best service, and quality" (Company H, Sales Manager).

These findings corroborate extant literature that emphasizes that co-branding strategies focus on providing benefits to all involved brands (e.g., Egan, 2011). Overall, this study demonstrates that being acknowledge by all partners as mutual beneficial is a critical success factor to co-branding strategy.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Partner Selection Criteria

As the leader of this co-branding partnership, the Hotel had a fundamental role in developing the strategy and, ultimately, it is responsible for choosing its wine partner brands. Hotel's representatives consistently identified product quality as the main criteria for partner selection. As mentioned by the Hotel's Brand Assistant (Interviewee 4) "being a luxury brand, the Hotel must be associated to similar brands that are perceived as premium, prestige brands by the consumer". Hence, initially it was essential for the Hotel to reinforce its brand positioning by the association with the most reputable wine brands, benefiting from the transfer of credibility. These findings confirm that the selection of co-branding partners must focus on brand positioning (Wason & Charlton, 2015) and performance (Geylani et al., 2008).

This study shows that as the Hotel brand gained its own prestige, it could collaborate with a different range of brands. Partners must provide high quality products, but the Hotel representatives recognize that nowadays they are able to cooperate with unknown brands – as long as their products present high quality standards, as explained in the following sections.

Co-Branding Impact on Brand Positioning

The fact of being a vinous-concept luxury hotel obviously favors the adoption of co-branding strategy with high-quality wine producers. Indeed, the Hotel's representatives recognize that the association with wine brands is essential to hotel's brand positioning. As mentioned by the Marketing CEO (Interviewee 6) co-branding "helped a lot to position the hotel as one of the best vinous-concept hotels in Europe. The diversity of partners adds a lot of value to this aspect".

The co-branding strategy that was undertaken by the Hotel was essential for its positioning, consolidating its vinous-concept and its role of wine embassy. Besides the benefits for the Hotel itself, this stronger positioning provides long-term benefits that can be also transferred to wine company partners. Indeed, regarding the impact on wine brands' positioning, the advantage of being associated with a luxury hotel was also highlighted. For instance, the Hotel's Brand Assistant (Interviewee 4) explained that

The [Hotel] quickly achieved a premium positioning, given the quality of its facilities, the service it provides, the quality of its restaurants, among many other factors. Being associated with [Hotel] means being associated with a brand of luxury and refinement, not accessible to everyone.

Overall, wine companies participating in this study consistently noted that the luxury vinous-concept positioning of the Hotel favours wine brands. The Marketing and Sales Manager of Company F states that "it is of utmost importance and reason of pride for our brand to be associated to a luxury hotel, because it shows that we have quality, image, and brand characteristics that allow such partnership". Company C interviewee further notes that the advantages for wine brand positioning is also due to the exclusivity of the partnership:

Being amongst a relatively small and highly selected number of partners, who produce very high quality wines, contributes to enhancing the credibility and confidence of consumers regarding our own wines (Marketing and Sales Manager, Company C).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thus, these findings offer empirical evidence that co-branding helps reinforcing brand positioning (Singh et al., 2014). The role of the stronger brands as endorsers of the quality of their partners is also evident, corroborating extant literature (e.g., Genc, 2010; Singh et al., 2014).

Co-Branding Impact on Luxury/Premium Brand Image

As demonstrated in previous sections, brand image is one of the most discussed consequences of co-branding strategies. Without surprise, participants in this study stress that co-branding strategy had an undeniable impact, particularly important for the weaker brands. In the words of the Hotel's CMO (Interviewee 6), "(...) the [Hotel] brand is stronger than most producers', so in many cases it was the producer who benefits most [from the partnership]". While it seems obvious that being associated to some well-known wine brands was essential for the vinous-concept Hotel's credibility and luxury positioning, the Hotel managers concluded it were the wine brands that took the most advantages. Mainly, because "for these brands undoubtedly it is an asset to be associated with us, especially regarding awareness" (Interviewee 4, Marketing Assistant). Aspect confirmed by several partners, as for instance the CEO of the wine Company G "[the partnership] strengthens our awareness, allowing us to reach a target of high-end customers that buy luxury products".

Hence, in this case co-branding strategy is essential to develop a premium image for wine brands, certifying the quality of their products. Interviewee 1 (Wine Director) referred that this applies even to very small wine brands "there are small producers that have never been detected by wine quality radars but being here it is almost a certification". Clearly, one can expect a co-branding strategy be particularly important for unknown brands and small companies, as suggested, for instance, by the Sales Manager of Company B: "for unknown brands it's very difficult to enter the market. Therefore, it was very important to us to be associated with a hotel with such a brand and charisma".

However, the rewards in terms of brand image were also acknowledge by the medium and large companies with stronger brands. As explained by the Key Account Manager of Company D, their brands had a strong and positive image before the partnership, but "in general all brands associated with [the Hotel] may benefit from a premium image". Marketing and Sales Manager of Company C, one of the large companies participating in this study proposes the reasons for the benefits provided even to well-known brands:

The Hotel has developed a strong premium image due to their target customers and the very high quality and variety of services it provides. Therefore, the brands that partner with it benefit from the association with its quality and high reputation (Marketing and Sales Manager, Company C).

Hence, consistently with literature (Lee, 2014; Park et al., 1986), findings highlight the co-branding ability to reinforce brand image. Although co-branding mutually reinforces the partner brands, as suggested by Geylani et al. (2008), the advantages seem more evident for the less known and reputable brands, as suggested by authors such as Abratt and Motlana (2002). This study also offers evidence that this process has a dynamic nature. In the beginning, apparently the Hotel brand most benefited from the association with the prestigious wine brands. Then, as Hotel's image consolidated over time, it assumed the role of reinforcing the brands of its partners, especially the brands of the smaller producers. Overall, this study corroborates Geylani et al. (2008) arguments that one of the main goals of co-branding is

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

reinforcing the image of partner brands and confirms that the stronger brands assume an important role as endorsers of their partners.

The Hotel as a Channel Intermediary

This co-branding strategy was referred by all participants as successful and with positive impact on the partner brands. However, the role of the partnership in terms of distribution channel did not showed up from interviews, except in the case of very small producers, such as Company E, whose CEO claimed that the Hotel was the main distribution channel for their wines. For the CEO of the medium-sized Company G, the Hotel's wine shop, bars and restaurants are "privileged distribution channels", and Company D Key Account Manager expects that "the Hotel can leverage sales inside and outside the national market". In the opinion of the Hotel's CMO "without this type of partnership it would be impossible for small producers to target this kind of customers". This corroborates Helmig et al. (2008) who claim that co-branding help reduce costs especially in the case of new products.

SOLUTIONS AND RECOMMENDATIONS

This study provides interesting cues for managers, more specifically in the luxury hospitality and wine business sectors, on how to use co-branding to reinforce brand positioning. Moreover, in contrast with most of previous research on the topic, research investigates co-branding in luxury hospitality management from the managers' point of view. Thus, it constitutes an interesting case for comparing co-branding management practices.

This study points out that co-branding involving companies from different sectors offers interesting opportunities for differentiation, positioning, and for developing a strong and positive brand image, especially in the case of involving producers that directly connect with the concept defined by the brand. Considering that this article aimed primarily at determining the benefits of co-branding, some aspects for managing partnership strategies between brands from different business sectors should be pointed out.

The literature notes that managers operating in the hotel business must join forces to achieve competitive advantage (Li et al., 2014) and maximize the value offered to the consumers (Su & Reynolds, 2016). Throughout this study, it was possible to verify that this partnership provides several benefits to the brands involved, including brand awareness and reach. Therefore, it is advisable that less-known wine brands that aspire to a prestigious reputation associate themselves with hospitality luxury brands, namely taking into account that the wine market is highly differentiated and fragmented (Brochado et al, 2015; Karelakis et al., 2008). Collaborating with high-end hospitality brands can help to leverage wine brands that wish to stand out, attract new customers, and achieve differentiation and competitive advantage. Therefore, wine producers looking for differentiation competitive advantage could cooperate with brands of different product categories, in order to enhance their products' distribution, communicate the quality of their brands, and reach other potentially interested audiences.

The study also provides recommendations for hotels and other companies that intend to develop a differentiated positioning through partnerships that are capable of adding value to their brands. The establishment of co-branding strategies with companies from different business sectors enhance the success of the brands involved by the ability to differentiate and position it (Hadjicharalambous, 2006; cit. in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Guillet & Tasci, 2012). In addition, co-branding presents interesting opportunities to improve brand image, as partner's brand reputation provides credibility, thus allowing effective brand image management.

FUTURE RESEARCH DIRECTIONS

It is important to refer that the study focused only on one ongoing co-branding initiative, with a limited number of partners actively involved in the study. Therefore, we recommend that future research further investigate this topic under different contexts and research settings, in order to validate findings. Suggestions for future research include comparing co-branding strategies in luxury hotels with different products and services. Studying unsuccessful co-branding strategies would also be useful to both academics and practitioners.

In addition, it would be interesting to investigate consumers' perceptions regarding co-branding strategies. One possible direction for future research is carrying out quantitative studies in order to validate some of the results achieved with qualitative data. Based on the literature review and on the results of this exploratory study, we formulated four hypotheses of study. These hypotheses represent a framework of analysis for future studies.

H1. Co-branding strategies between brands of different business areas influence positively consumers' perceptions about the brands involved in the partnership.

According to Abratt and Motlana (2002), associations created by co-branding strategies can both increase brand value and enhance or damage consumers' perceptions of each of the brands involved. In fact, there may be a lack of control in the alignment of both brands in the minds of consumers (Helmig et al., 2008). Several factors can influence consumers' perceptions on branding partnerships, such as, disappointment on brand performance that may result from exacerbated consumer expectations, and the impact of attitudes toward brands prior to the partnership on their partner brands' associations and evaluations (Abratt & Motlana, 2002).

H2. Brand reputation and perceived quality have a positive impact on partner's brand reach and image.

Co-branding allows the exchange of benefits in terms of brands' perceived quality and reputation (Cegarra & Michel, 2001), and potentially adds value to product brands (Jakubanecs & Supphellen, 2012). Co-branding is one of the techniques that has been increasingly used to reinforce brand image (Kottemann et al., 2017; Lee, 2014; Warraich et al., 2014). In addition, results of this study suggest that product brands benefit more with the partnership with the luxury hotel, because the recognition in the market and the strength of its positioning. Consequently, product (wine) brands can be perceived as prestige brands due to this partnership. Thus, the partnership allows product brands to be associated with luxury and high quality, consolidating and boosting their brand image.

H2.1. Partnership with a luxury hotel brand improves positively an image of prestige for the product brands.

According to our findings, the luxury positioning of the Hotel brand, the high standards of the service quality it provides, and its awareness create value and favour the development of a prestige image for the wine brands.

This study shows that the luxury hotel acts as an endorser and supports the promotion of the product brand and its offers, providing empirical support to contributions in the literature (e.g., Hendrasaputra & Lunarjanto, 2009). Service providers have several opportunities to feature partner brands, such as through the organization of events, the customization of hotel facilities, and the sale of products through

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the hotel's stores. Through these activities, it is possible to reach potential consumers and disseminate an image of quality of the wine brands. Both well-known wine brands and newly launched brands recognise these benefits.

H3. The credibility of one brand positively influences the image of the brands in the partnership.

The endorser grants support and credibility to the partners. It guarantees that it will be able to provide all the desired functional benefits (Aaker, 1996). In the wine sector, the association with other brands results in numerous advantages, including high credibility and perceived quality (Vrontis et al., 2010). In addition, according to this study, the luxury brand benefited from the quality and reputation of some wine partners in the sector, which provided the brand with credibility in the wine sector. Likewise, the quality and prestige of the hotel granted credibility and quality certification to the product brands.

CONCLUSION

Considering that literature on co-branding is still scarce and has disregarded the luxury hospitality sector, this article provides interesting contributions on the topic while confirms the relevance of co-branding strategies between hotels and products directly connected to its concept, positioning or differentiating characteristics. The empirical results shared in this article demonstrate the characteristics and impact of a successful co-branding strategy implemented by a vinous-concept luxury hotel and high-quality wine brands, as described by Hotel and Wine Companies' managers. Helmig et al. (2008) emphasize that commitment of all partners is essential for co-branding success. Commitment is easily achieved when all partners get and are aware of clear benefits, in terms of not only profitability, but also regarding brand image and positioning. Results suggest the improvement of brand image as one of the main advantages of co-branding with a luxury hotel. Indeed, all participants in this study acknowledge that this co-branding strategy has a positive and enduring impact on brand image and reinforces brand positioning. Nevertheless, participants also emphasize that the coherence of the brand values and the positioning of the partners involved in the co-branding are essential. More specifically, in the case of a co-branding involving prestige brands, the level of the quality of the products involved is crucial. Therefore, despite the advantages that a co-branding with a strong brand may represent for brands still fighting for awareness and positioning, this partnership is particularly demanding for small size partners. As noted by one of the participants in this study "the true partnership forces us to be in our markets as good as they are in theirs" (Company E, CEO).

REFERENCES

Aaker, D. A. (1996). *Building Strong Brands*. Simon & Schuster.

Abratt, R., & Motlana, P. (2002). Managing co-branding strategies: Global brands into local markets. *Business Horizons*, *45*(5), 43–50. doi:10.1016/S0007-6813(02)00242-2

Balmer, J. M. T., & Greyser, S. A. (2002). Managing the multiple identities of the corporation. *California Management Review, 43*(3), 72-86. doi:10.2307/41166133

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Boone, J. M. (1997). Hotel-restaurant co-branding: A preliminary study. *The Cornell Hotel and Restaurant Administration Quarterly*, *38*(5), 34–43. doi:10.1016/S0010-8804(97)86581-0

Brennan, R., Canning, L., & McDowell, R. (2014). *Business-to-Business Marketing* (3rd ed.). Sage.

Chiambaretto, P., & Gurău, C. (2017). David by Goliath: What is co-branding and what is in it for SMEs. *International Journal of Entrepreneurship and Small Business*, *31*(1), 103–122. doi:10.1504/IJESB.2017.083805

Dahlstrom, R., & Dato-on, M. C. (2004). Business-to-business antecedents to retail co-branding. *Journal of Business-To-Business Marketing*, *11*(3), 1–22. doi:10.1300/J033v11n03_01

Dev, C. S., Morgan, M. S., & Shoemaker, S. (1995). A positioning analysis of hotel brands: Based on travel-manager perceptions. *The Cornell Hotel and Restaurant Administration Quarterly*, *36*(6), 48–55. doi:10.1177/001088049503600617

Díaz-Bustamante, M., Carcelén, S., & Puelles, M. (2016). Image of luxury brands: A question of style and personality. *SAGE Open*, *6*(2), 1–15. doi:10.1177/2158244016644946

Egan, J. (2011). *Relationship Marketing: Exploring Relational Strategies in Marketing*. Pearson Education.

Erevelles, S., Stevenson, T. H., Srinivasan, S., & Fukawa, N. (2008). An analysis of B2B ingredient co-branding relationships. *Industrial Marketing Management*, *37*(8), 940–952. doi:10.1016/j.indmarman.2007.07.002

Genc, R. (2010). Strategic brand management in hospitality sector: How to manage co-branding in hotels and restaurants. *Acta Universitatis Danubius Œconomica*, *6*(3), 33–46.

Geylani, T., Inman, J. J., & Hofstede, F. T. (2008). Image reinforcement or impairment: The effects of co-branding on attribute uncertainty. *Marketing Science*, *27*(4), 730–744. doi:10.1287/mksc.1070.0326

Glaser, B. G., & Strauss, A. L. (2006). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Transactions.

Goulding, C. (1999). Grounded theory, ethnography and phenomenology: A comparative analysis of three qualitative strategies for marketing research. *European Journal of Marketing*, *39*(3/4), 294–308. doi:10.1108/03090560510581782

Grossman, R. P. (1997). Co-branding in advertising: Developing effective associations. *Journal of Product and Brand Management*, *6*(3), 191–201. doi:10.1108/10610429710175709

Guillet, B. D., & Tasci, A. D. (2012). Chinese hoteliers' take on hotel co-branding in China. *Tourism Review*, *67*(4), 3–11. doi:10.1108/16605371211277777

Helmig, B., Huber, J.-A., & Leeflang, P. S. (2008). Co-branding: The state of the art. *Schmalenbach Business Review*, *60*(4), 359–377. doi:10.1007/BF03396775

Kapferer, J.-N. (1998). Why are we seduced by luxury brands? *Journal of Brand Management*, *6*(1), 44–49. doi:10.1057/bm.1998.43

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kapferer, J.-N. (2012). *The New Strategic Brand Management: Advanced Insights and Strategic Thinking* (5th ed.). Kogan Page.

Keller, K. L., & Richey, K. (2006). The importance of corporate brand personality traits to a successful 21st century business. *Journal of Brand Management*, *14*(1-2), 74–81. doi:10.1057/palgrave.bm.2550055

Lee, C.-L. (2014). Is co-branding a double-edged sword for brand partners? *European Research Studies Journal*, *17*(4), 19–34. doi:10.35808/ersj/430

Lewis, R. C. (1981). The positioning statement for hotels. *The Cornell Hotel and Restaurant Administration Quarterly*, *22*(1), 51–61. doi:10.1177/001088048102200111

Li, X., Yen, C.-L., & Uysal, M. (2014). Differentiating with brand personality in economy hotel segment. *Journal of Vacation Marketing*, *20*(4), 323–333. doi:10.1177/1356766714527965

Malik, M. E., Naeem, B., & Munawar, M. (2012). Brand image: Past, present and future. *Journal of Basic and Applied Scientific Research*, *2*(12), 13069–13075.

Moon, H., & Sprott, D. E. (2016). Ingredient branding for a luxury brand: The role of brand and product fit. *Journal of Business Research*, *69*(12), 5768–5774. doi:10.1016/j.jbusres.2016.04.173

Motion, J., Leitch, S., & Brodie, R. J. (2003). Equity in corporate co-branding: The case of Adidas and the All Blacks. *European Journal of Marketing*, *37*(7/8), 1080–1094. doi:10.1108/03090560310477672

Nobre, H. (2019). A marca como ferramenta de gestão empresarial. In C. Machado & J. P. Davim (Eds.), *Organização e Políticas Empresariais* (pp. 133–166). Actual Editora.

Nobre, H., & Simões, C. (2019). NewLux brand relationship scale: Capturing the scope of mass-consumed luxury brand relationships. *Journal of Business Research*, *102*, 328–338. doi:10.1016/j.jbusres.2019.01.047

Nueno, J. L., & Quelch, J. A. (1998). The mass marketing of luxury. *Business Horizons*, *41*(6), 61–61. doi:10.1016/S0007-6813(98)90023-4

Öberg, C. (2016). What creates a collaboration-level identity? *Journal of Business Research*, *69*(9), 3220–3230. doi:10.1016/j.jbusres.2016.02.027

Okonkwo, U. (2009). The luxury brand strategy challenge. *Journal of Brand Management*, *16*(5-6), 287–289. doi:10.1057/bm.2008.53

Park, C. W., Jaworski, B. J., & MacInnis, D. J. (1986). Strategic brand concept-image management. *Journal of Marketing*, *50*(4), 135–145. doi:10.1177/002224298605000401

Riley, D., Charlton, N., & Wason, H. (2015). The impact of brand image fit on attitude towards a brand alliance. *Management & Marketing*, *10*(4), 270–283. doi:10.1515/mmcks-2015-0018

Shen, B., Choi, T.-M., & Chow, P.-S. (2017). Brand loyalties in designer luxury and fast fashion co-branding alliances. *Journal of Business Research*, *81*, 173–180. doi:10.1016/j.jbusres.2017.06.017

Singh, J. P., Kalafatis, S., & Ledden, L. (2014). Consumer perceptions of cobrands: The role of brand positioning strategies. *Marketing Intelligence & Planning*, *32*(2), 145–159. doi:10.1108/MIP-03-2013-0055

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Su, N., & Reynolds, D. (2017). Effects of brand personality dimensions on consumers' perceived self-image congruity and functional congruity with hotel brands. *International Journal of Hospitality Management*, *66*, 1–12. doi:10.1016/j.ijhm.2017.06.006

Tynan, C., McKechnie, S., & Chhuon, C. (2010). Co-creating value for luxury brands. *Journal of Business Research*, *63*(11), 1156–1163. doi:10.1016/j.jbusres.2009.10.012

Vrontis, D., Thrassou, A., & Czinkota, M. R. (2011). Wine marketing: A framework for consumer-centred planning. *Journal of Brand Management*, *18*(4-5), 245–263. doi:10.1057/bm.2010.39

Wang, S.-C., Soesilo, P. K., & Zhang, D. (2015). Impact of luxury brand retailer co-branding strategy on potential customers: A cross-cultural study. *Journal of International Consumer Marketing*, *27*(3), 237–252. doi:10.1080/08961530.2014.970320

Wason, H., & Charlton, N. (2015). How positioning strategies affect co-branding outcomes. *Cogent Business & Management*, *2*(1), 1092192. doi:10.1080/23311975.2015.1092192

ADDITIONAL READING

Beverland, M. (2018). *Brand Management: Co-creating Meaningful Brands*. Sage.

Heding, T., Knudtzen, C. F., & Bjerre, M. (2020). *Brand Management: Mastering Research, Theory and Practice*. Routledge. doi:10.4324/9780367172596

Kapferer, J. N. (2012). *The New Strategic Brand Management: Advanced Insights and Strategic Thinking*. Kogan Page.

Kapferer, J. N., & Bastien, V. (2012). *The Luxury Strategy: Break The Rules of Marketing to Build Luxury Brands* (2nd ed.). Kogan Page.

Keller, K. L., & Swaminathan, V. (2019). *Strategic Brand Management: Building, Measuring, and Managing Brand Equity* (5th ed.). Pearson Education.

Linder, C., & Sperber, S. C. (2018). Leveraging reputation for product innovation through strategic co-branding. In Information Resources Management Association (Eds.) Driving Customer Appeal Through the Use of Emotional Branding (pp. 248-261). Hershey: IGI Global. doi:10.4018/978-1-5225-2921-7.ch014

Rosenbaum-Elliott, R., Percy, L., & Pervan, S. (2015). Strategic Brand Management (4th ed.). Oxford University Press.

Washburn, J. H., Till, B. D., & Priluck, R. (2000). Co-branding: Brand equity and trial effects. *Journal of Consumer Marketing*, *17*(7), 591–604. doi:10.1108/07363760010357796

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Brand Equity: The value that a brand provides to its products and services.

Brand Image: Customers' associations and beliefs about a brand.

Brand Perceptions: Positive or negative feelings that result from the consumers' experiences with the brand.

Brand Positioning: A strategy which main aim is to make the brand differentiated from the competitors' brands in the mind of the consumers.

Brand Value: It is often a synonym of brand equity.

Co-Branding: A strategy that creates associations between two or more brands of strategic partners, namely in the form of brand partnerships or brand collaborations.

Customer Loyalty: A relationship between customer and brand that results in repeated purchases and often in the preference of that brand instead of its competitors.

Spillover Effects: An economic event that results from another apparently unrelated event.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 18
Data Analysis in the Shipping Industry:
eShip Case Study – Problem Statement

Marcel Kyas
Reykjavik University, Iceland

Joshua D. Springer
Reykjavik University, Iceland

Jan Tore Pedersen
Marlo as., Norway

Valentina Chkoniya
https://orcid.org/0000-0003-1174-3216
University of Aveiro, Portugal

ABSTRACT

This chapter identifies the critical issues that must be addressed to accelerate the digital transition in the chartering market. The maritime industry is one of the pillars of global trade, where change is a constant. Again, shipping is at the cusp of a new era—one driven by data. The authors review the state-of-the-art technology that is useful to automate chartering processes. · The Fourth Industrial Revolution (or Industry 4.0) starts to change the bulk shipping markets leveraging the data flow between industrial processes in the physical and virtual world. · The internet of things accelerates data flow from things in the real world to the virtual world and enables us to control processes in real-time. Machine-to-machine communication, together with artificial intelligence, creates autonomous systems in many areas of production and logistics. Based on the gathered elements, eShip's case study was analyzed, and future steps have been defined for the data analysis in the shipping industry.

DOI: 10.4018/978-1-7998-6985-6.ch018

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Traditionally, the shipping industry has been highly dependent on its various mini-ecosystems, consisting of ports, authorities, ships, ship operators, shipowners, cargo owners, and many more. However, over a decade ago, researchers noted that the traditionally conservative shipping industry is undergoing a change, where it is believed that the demands for an increase in efficiency, safety, and protection of the environment can be only achieved by more innovation (Jacks & Pendakur, 2010; Perunovic & Vidic, 2011; SmartPort, 2019). Half a century ago, the French geographer Perpillou considered ports and maritime transport to form a constellation (Ducruet & Zaidi, 2012). All the different players in the ecosystem are inextricably linked, making interconnectivity and cooperation between them crucial for effective and efficient operations of the transport system (Martimo, 2017). The port of the future is a smart port, and the ship of the future is a smart ship. However, before getting smart, all the "constellations" must go digital (Delenclos, Rasmussen, & Riedl, 2018). Change is a constant in the maritime industry, and it is once again at the cusp of a new era—one driven by increased digitalization and innovation (Levinson, 2020; SmartPort, 2019). Smart ships and smart ports will exchange data for speeding up the processes at the port, including refueling, unloading, and loading cargo. Paperwork can be handled automatically. Ports will exchange data, announcing what leaves one port to the ports that receive the cargo. Such exchange speeds up customs clearance.

The fourth industrial revolution promises increased automation of all processes. This includes the shipping industry. Parcel services like DHL track their shipments around the world in real-time (Diwan, 2016). This development necessitates that all stations the cargo takes are digitized. For the shipping industry, this includes the ships and the ports. The Boston Consulting Group identified several emerging technologies like advanced analytics, autonomous shipping, robotics, and artificial intelligence set to change how planning, operations, commercial, and support functions within shipping are performed (Delenclos et al., 2018). With the addition of digital data about cargo and ships, data analytics and machine learning uncover better ways to handle ships, ports, and cargo. It enables to measure the impact of every step, decrease cost, and speed up processes.

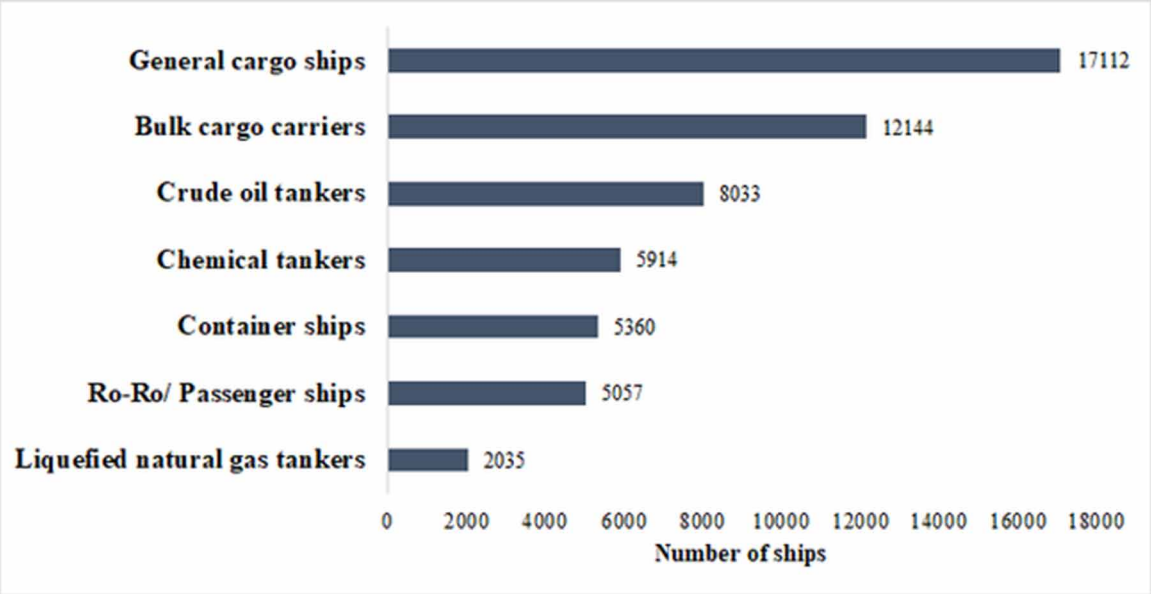
Information technology on-board and in communication from vessels to ports and authorities are developing as required. However, according to a survey conducted by eShip in 2021, the dialogue between cargo owners (charterers) and shipping operators is, in most cases, still using phones and various forms of paper, spreadsheets, and other unstructured data formats for keeping and sharing information (PDF documents in e-mails being one example). Consequently, there is significant room for improvement, particularly when organizing operations and making the interaction between stakeholders more efficient. The authors review the state-of-the-art technology useful to automate chartering processes that serve as a base for the eShip project development. The eShip project, funded by an EEA Grant, is developing an online platform for making the entire chartering process digital. This includes contracts of affreightment, charter party terms, real-time voyage tracking, demurrage/dispatch calculations, freight invoice generation, bills of lading, post-voyage analysis, and regulatory compliance. Based on the gathered insights, the critical issues that must be addressed to accelerate the digital transition in the chartering market were identified. Future steps have been defined for the data analysis in the shipping industry.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

The maritime industry is one of the critical pillars of global trade (Cacho, Marques, & Nascimento, 2020; Chkoniya, Madsen, & Bukhrashvili, 2020). According to the International Chamber of shipping, some 11 billion tons of goods are transported by ship each year. This represents an impressive 1.5 tons per person based on the current global population. Shipping's capacity to transfer goods and materials from where they are produced to where they will be consumed underpins modern life. Of this, 2 billion tons is crude oil, 1 billion tons is iron ore (the raw material needed to create steel), and 350 million tons of grain. Such and other similar shipments would not be possible by road, rail, or air. Large volumes of other vital products such as chemicals, refined fuels, and manufactured goods are also shipped by sea. The world fleet of commercial ships as of the beginning of 2020 (Statista 2021) is shown in Figure 1.

Figure 1. World fleet of commercial ships (adapted from Statista, 2021)



The cargo shipping industry transports, broadly, six kinds of cargo: bulk solids, bulk liquids, containerized units, refrigerated/chilled, roll-on, roll-off cargo, and general. General cargo includes everything not covered by the previous kinds of cargo. Research in the handling of cargo has overwhelmingly focused on containerized cargo and palletized cargo. These forms of cargo show many advantages over bulk cargo. Handling is standardized, and the cargo can shift from one mode to another, e.g., loaded from a ship to a train or truck without effort.

Looking at the shipping industry, container shipping, an integral part of logistics operations dealing with fast-moving consumer goods, has been forced to employ information technologies for booking, recording loading of containers, and reporting the status of container movements (Haralambides, 2019). It has also been required to connect electronically to those dealing with hinterland transportation. Specialized shipping operations, like the autonomous ship Yara Birkeland and innovative inland waterway

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

operations like the ones managed by Seafar in Antwerp (where the captain of vessels sits in a control room on land) demonstrate that information technology is extensively used in maritime transportation.

The publication of the three books on Maritime Informatics also shows that the topic is gaining significant academic and commercial attention (Lind, Michaelides, Ward, & Watson, 2020). However, not all shipping areas have been equally fortunate or forward-looking to adopt innovative technologies for operational purposes. Much of the maritime sector has adopted the habit of adopting innovative technologies only after being forced to it by authorities (Ward & Bjørn-Andersen, 2021).

Shipping operations dealing with wet and dry bulk cargo have not faced the same pressures from their surrounding stakeholders that are the case for container shipping (Mileski, Clott, Galvao, & Laverne, 2020). Information technology on-board and in communication from vessels to ports and authorities are developing as required. However, the dialogue between cargo owners (charterers) and shipping operators is in most cases still using phones and various forms of paper, spreadsheets, and other unstructured data formats for keeping and sharing information (PDF documents in e-mails being one example). Consequently, there is significant room for improvement, particularly when organizing operations and making the interaction between stakeholders more efficient.

Research in Shipping

Research in shipping started with the founding of Operational Research at Brawdsey Research Station by A.P. Rowe. Before this, the shipping industry lacked a scientific approach. Mathematics was recast in a modern way, and computers were under development. The first problems of Operational Research and shipping were the logistics of the second world war. Troops needed to be moved and then supplied. Ships were the primary method of transport (Gass & Assad, 2005).

Then, research focused on shipping containerized cargo. Especially, containers offered many exciting problems. This resulted in an abundance of publications on container shipping.

At the same time, the problems of the bulk shipping market have received little attention. With bulk shipping, we focus on dry bulk, liquid bulk, and breakbulk shipping: dry bulk includes ores and grains, liquid bulk includes petrochemicals and liquidized natural gas, and breakbulk includes bagged coffee, barrels, metal drums, automobiles, and machine parts.

Chartering is the process of setting up an agreement (Charter Party) between a charterer (the one who owns the cargo and needs to ship it) and shipowners/operators for the cargo carriage. It may seem a straightforward process, but typically involves many stakeholders, including brokers that typically match the demand for transport with the ships with the appropriate capabilities to move cargo. Involving many parties, the process may be pretty complex. Information systems for supporting the chartering process have been around for quite a long time. Companies like ShipNet, Inchcape, Danaos, SeaSoft, etc., have been supplying such solutions for decades. One question then arises: Since these systems have been around for such a long time, why are they not extensively used? Why is the chartering process manual in most operations?

There may be many answers, but one is predominant: cost of implementation and use. These early systems required to be implemented in the users' premises, initial implementation and licenses were expensive, and software maintenance costs were typically 15% of the initial license per year. The operation needed to be significant to afford such investments. The introduction of cloud technologies creates a new change in thinking in this business. There is no longer the need for on-site implementation. New software-as-a-service business models make the use of such systems much more affordable and flexible.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The early providers of such information systems aim to "place" existing systems in the cloud, creating a "new layer" on top of them. However, if we put a new bridge on an old ship, we are making some things more accessible, but we are not changing the essential capabilities of the ship—the same thing with the information system. To make the most of cloud solutions, we need to design the systems for the cloud.

This is what eShip is doing. When these new, low-cost solutions are being used, data from operations will be stored and then available for further analysis, which should benefit all aspects of the chartering process.

Shipping 4.0

Accordingly, the term Shipping 4.0 was coined in 2016 to describe the new developments in the digitalization of shipping, to reflect the similar developments in the land-based industry, which commonly goes under Industry 4.0 (Kavallieratos, Diamantopoulou, & Katsikas, 2020). The Fourth Industrial Revolution (or Industry 4.0) is automating manufacturing and industrial practices using smart technology. Factories integrate large-scale machine-to-machine communication (M2M) and the Internet of Things (IoT). They use improved communication and self-monitoring and produce smart machines that can analyze and diagnose issues without human intervention. The term "Industry 4.0" describes different – primarily IT-driven – changes in manufacturing systems. These developments do not only have technological but furthermore versatile organizational implications. As a result, a change from product- to service orientation even in traditional industries is expected. Second, an appearance of new types of enterprises can be anticipated which adopt new specific roles within the manufacturing process responsible for the value-creation networks (Scheer, 2012). For instance, it is possible that comparable to brokers and clearing-points in the branch of financial services, analogue types of enterprises will also appear within the industry. With the planning, analysis, modelling, design, implementation, and the maintenance (in short: the development) of such complex, dynamic, and integrated information systems, an attractive and at the same time challenging task for the academic discipline of BISE (Business & Information Systems Engineering) arises, which can secure and further develop the competitiveness of industrial enterprises (Lasi, Fettke, Kemper, Feld & Hoffmann, 2014). The Fourth Industrial Revolution is thus expected to have a disruptive impact on maritime and shipping sectors, where smart ships and smart ports will be part of a new and fully interconnected maritime ecosystem. In such a context, two important questions arise about what the current level of technological advancement in the maritime shipping industry is and how the new technologies and methodologies should be deployed to ensure a rapid transition to Shipping 4.0 (Aiello, Giallanza, & Mascarella, 2020).

MAIN FOCUS OF THE CHAPTER

Unlike container shipping that is entirely digitized, bulk shipping, which represents 2/3 of the global maritime transport volume, is in dire need of digital transformation as it still relies on non-structured information exchange. However, bulk shipping is now beginning to adopt a smarter approach to how it best goes about its transformation towards a global value chain. In this case study, we analyze eShip, an online platform that applies state-of-the-art information and communication technology (ICT) tools to the traditional chartering business model, generating significant value added to all market participants in three key aspects.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

eShip Project

The eShip project is funded under EEA Grants, a multiannual financial mechanism sponsored by Iceland, Liechtenstein, and Norway. eShip was ranked #1 among numerous candidates to the EEA Grants Call 1.023.2019. Its objective is to develop a cloud-based (also called Software-as-a-Service -SaaS) chartering application to support the traditional chartering process in bulk shipping - digitizing and automating the entire value chain.

As indicated before, many participants are involved in establishing a charter party. Consequently, communication between these parties is a significant challenge, particularly when trying to standardize information that has not yet resulted in intangible results. Already Lord Kelvin wrote that to know something, one must measure it, which is often restated as "to control," "to manage," "to optimize," and more. Measurements generate troves of data, which need to be interpreted. With data analytics, the authors reference all methods that help to interpret the data, for example, to understand how one measured quantity influences others. Understanding such relations then allows optimizing operations.

The members of the eShip project surveyed potential stakeholders. Data was collected in June 2021. Twenty-five responded (which is sizeable, considering the small size of the business). The majority consider digital transition important to the organization they represent (76%), identifying the need for digital platforms to support their chartering enterprises. An essential feature of such a system is to integrate with existing ERP systems (75%). The stakeholders want to leverage the data generated on ships, ports, and the chartering business to optimize their operations, minimize their impact on the environment, and demonstrate compliance with laws and standards.

The absence of Data Analysis is the main shortcoming they face (46%), followed by too many manual inputs and too many communication channels (38% each). Lack of synchronization between procedures and out-of-date functionality was mentioned by 25% of respondents. The results point that by providing a digital platform that may be used by all parties involved, either through a web interface or through EDI/PI interfaces, we will be able to maximize interoperability between all the participants, reuse information and reduce management costs and accelerate the digital transition in what is a complex and traditionally conservative sector.

eShip will contribute to this end in 3 key areas:

1. Automation of information-heavy procedures during the pre-fixture and post-fixture stages of the chartering process.
2. Seamless integration with existing management systems, particularly ERPs, transport management systems, and port community systems.
3. Application of business intelligence and data analytics to optimize the entire value chain and support the rapid transition to green shipping.

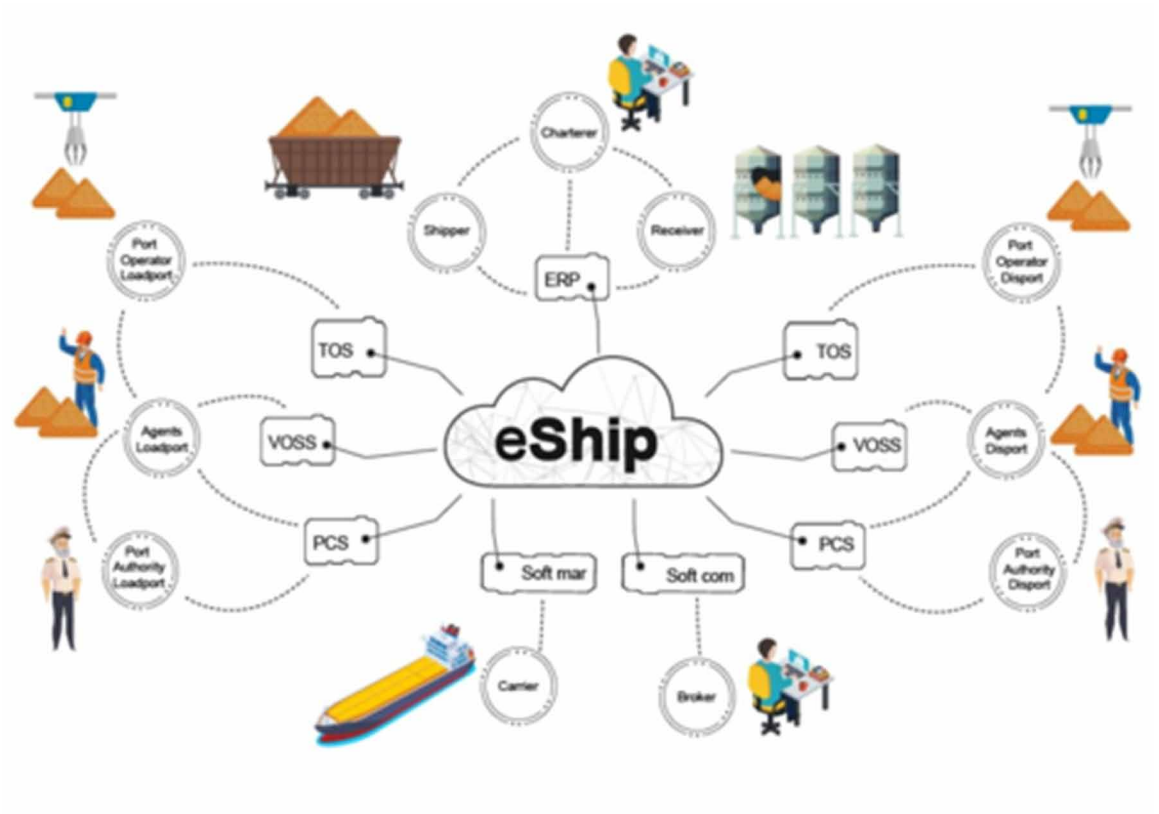
The eShip project provides a cloud platform for the entire ship chartering process in an end-to-end capacity (Figure 2). It includes the following capabilities:

- Search for a ship function in the charterer market and the subsequent freight negotiation and contract agreement.
- Monitoring all the interlinking stages of the voyage – including tracking of the various voyage execution processes in real-time.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Recording all information related to billing (Bills of Lading and freight invoices) and invoicing calculations, including Statements of Facts and Laytime calculations.

Figure 2. eShip Ecosystem (©2021, eShip project, Used with permission)



The key benefits of using the SaaS solution developed in the eShip project will be:

1. Optimization of freight rates via intelligent matching of ships and cargo using smart algorithms.
2. Reduction of back-office expenses due to automation of repetitive and time-consuming procedures.
3. Open collaboration between participants and business units without disclosing information to unauthorized users.
4. Application of business intelligence and structured data analysis to achieve chartering process optimization.

An essential extra benefit arises from the choice of taking the Software-as-a-Service approach. To apply the solution, users need access to a web browser and the Internet. Initial licenses and implementation costs will be negligible. A typical SaaS business model is "pay per use," either as a fixed subscription fee per month or a fee based on transactions. With this, there should be no financial barriers for shall and hard-pressed companies to start using the platform.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

"Cargo is King" is a trading term frequently used in transport and logistics. Hence, the critical initial focus of the platform developed in eShip will secure benefits to the charterers. Those are the ones that will promote such solutions to be used by agents and shipowners/operators. However, to motivate all chartering process participants to use the solution, the needs of agents and other maritime stakeholders will also be providers. To be successful, the new digital platform needs to provide benefits for all users, reflected in the eShip project philosophy "Create value, Share value."

Digital Transition

The digital transformation of the shipping industry can enhance value within the industry's ecosystem (Poulis, Galanakis, Triantafillou, & Poulis, 2020). Emerging technologies are key change drivers in the shipping industry, despite their conservative nature (Table 1). As indicated in the market assessment report, the authors summarize the areas in which digital transition can reduce costs and improve performance from the shipowner's perspective, as recently expressed in the Future of Shipping webinar (Tat et al., 2020).

Single Windows

Singapore provides a good example where the Immigration and Checkpoints Authority (ICA), the Maritime and Port Authority of Singapore (MPA), and the Health Administration (ADA) now all combine data in a single submission. These organizations need to see even more combinations of data with shared governance - e.g., a container ship on the Asia / Europe trade via Singapore. Instead of the vessel reporting isolated data to Singapore, have the export data for China also become the import data for Singapore (with minor amendments). This process avoids individual submission to a single country (Singapore). The ADA reports that it is saving 200,000 person-hours per year by such measures. Authorities and ports will save even more if the data reporting occurs on the global level (cf. Portugal's JUL / JUP).

Lambrou, Watanabe, and Iida (2019) posit that digital transformation is a topical theme in shipping research and professional practice. They addressed technology and management aspects in a coalesced framework. They present five cases of advanced shipping incumbents' digitalization activities. Their theoretical model systematizes the technological components (technology typology), the prevailing management rationales (strategic drivers), and determinant factors (practices) of shipping digitalization. The authors of this report agree that the most critical technologies to further a digital transition are using artificial intelligence (AI) and data to optimize processes. IoT will play a crucial role in operating ships autonomously, digitizing commercial and business operations.

Real-Time Ship Tracking

One common question in the shipping industry is the estimated time of arrival. With a reasonable estimate, the charterer can order people and equipment to load and unload vessels, arrange transportation to or from the harbor, or arrange for interim storage. The authors see much use of the automatic identification system (AIS). AIS has become mandatory for large vessels (over 300 t) and allows tracking ships on the Internet. Ships may not send the correct data or send it irregularly, and receiving the signal depends on proximity to coastal base stations. Satellite communication is an option that some systems use.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Emerging technologies in the shipping industry

Tool	Description
Link Strategy to Digitization	Establish a core business strategy that dovetails with digitization ambitions to converge with the future strategy that the company wants to pursue.
Ship Data	Draw data from ships to build user cases with shipowners and ship managers.
Optimization	Better route planning & safety measures, e.g., bunker economics via route planning.
Ship-to-Shore	Indicators of Compromise Service (IOC) sensors for ship-to-shore connection.
Drone delivery	Spare parts delivery to the ship by drone (subject to current weight constraints). Ship and harbor inspection.
Transparency	Price discovery/offerings/settings for carriers, especially websites or platforms.
Remote Operation	Inspection / Surveys / Audits; the ship no longer waits for the documentation.
Classification societies	Digital class certificate replaces the traditional paper version, thereby avoiding the need for a class surveyor to deliver this to the ship.
Crewless vessels	promoting the need to transition towards new legal and navigational structures (a trend likely to gain momentum and further accelerate in a post-pandemic world).
Seafarer Mentality	New complexity surrounding crew changes: mental health challenges / increased self-harm are new features surrounding human emotions, spurred by Covid-19.
Data Lakes	The more data is collected, the greater the ability to build "data lakes" from simple origins initially and then build up to the sophisticated level, allowing the data to speak more clearly to you over time, thus allowing for improvements.
Chatbot	Instead of performing heavy-duty analytics, ask the specific query regarding the sought-after data point (e.g., operative speed/bunker consumption). Move away from an ambition to be very clever, concentrate on what is done with the available data - i.e., use it as the need arises.
Data platforms	What are these used for? What is the business case? They can be used as a vehicle for procurement processes and an aggregator for buyers and sellers in more efficient structures.
Analysis	Performance-related data from platforms. Noticing that means a rating, ranking, quotation, discussion, or analysis regarding an aspect of the investment performance of the platform, an asset allocation service, security, an index, or a benchmark.
Ecosystem Integration	Platforms allow for greater collaboration between shipyards, shipowners, AEMs, regulatory bodies, ship managers, engine manufacturers = everybody in the ecosystem, as compared to single sourcing: a collaborative solution is more efficient. Mass adoption leads to standardization; common usage allows for the digital off-take to be shared.
Standardization	It is a vast emerging trend; join forces to unify the approval process together.
Port community	Security, data exchange, platforms' interaction is all undergoing rapid re-sets.
Single Windows for Logistics	Facilitates trading and communication between economic operators and government agencies in the logistics sector. It offers a single administrative point of access to electronic exchanges related to international trade flows.

Companies like Marine Traffic offer application programming interfaces (APIs) to track ships in real-time and to query historical position data.

The data can use to track the ship and estimate its time of arrival at a port. The necessary equipment and crew to unload the ship can be ordered on time. The data can be analyzed for the seasonality of voyage time. Graziano, Renga, and Moccia (2019) describe a method of integrating AIS with radar images to estimate ship velocity and improve maritime situational awareness. The goal of their work is to avoid collisions of ships and delays in shipping. Chen, Khoo, Chong, and Yin (2014) demonstrate data-fusion to track ships and identify potential threats along the routes, including collisions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Use of Autonomous Vehicles and Vessels

Seaport Infrastructure Inspection: Unmanned air and sea vehicles can be helpful in infrastructure inspection in seaports, reducing financial and time costs and the number of areas people must inspect. Aerial drones are better able than humans to inspect hard-to-reach areas such as the top of a crane or shipping container. They can provide high-quality video records of the inspection. Sea-based vehicles can inspect underwater infrastructures such as ship hulls or supports. However, they are inherently limited by the water's visibility, making them applicable to only specific ports. On the other hand, they can use sonar or other technologies which enable data collection even in murky water. Overall, non-autonomous remotely controlled vehicles increase infrastructure inspection quality in port environments while requiring less time, cost, and risk (Stein, 2020).

Seaport Security: Unmanned vehicles can assist in seaports in compliance with the "International Ship and Port Facility Code" (Stein, 2018). Such crewless vehicles are to provide new points of view cheaply and without risk to personnel. Such applications include monitoring ships, cargo, traffic control, and emergencies. The crewless vehicles are typically remotely operated and serve as an extension of a human controlling the task. However, there is an opportunity for full automatization in cases of regular and unchanging tasks (such as the monitoring of docked ships and cargo).

Fully Autonomous Shipping: Autonomous ships provide rich potential for the shipping industry to cut costs while increasing operational times, shipping capacity, and reducing crew duties (Benson, D S & Colling, 2018). One such project is a Norwegian electric container ship called YARA Birkeland (Cross, Meadow, Com & Jostaphot, 2017), with a capacity of 120 twenty-foot equivalent units (TEU) (Autonomous ship project, critical facts about YARA Birkeland, 2017). The shipowner expects the YARA to launch in 2021 and sail among Herøya, Brevik, and Larvik, first with a small crew and then fully autonomously. Rolls-Royce (the Advanced Autonomous Waterborne Applications (AAWA) initiative) and Nippon Yusen Kabushiki Kaisha have undertaken similar projects. The latter has successfully tested an autonomous ship in Japan between the ports of Nagoya and Yokohama. The Maritime Unmanned Navigation through Intelligence in Networks (MUNIN) project has developed sensors for anticipated autonomous bulk carriers. It proposes that developing and using autonomous shipping vessels is technically feasible and could save financial and time costs (MUNIN, 2016).

One main challenge in autonomous shipping is developing a legal framework since the concept is so new to the legal world. For example, legal definitions of "autonomous ships" are not standardized or uniformly addressed – if addressed at all – from country to country. Similar is the case of a crew member's definition or remote operator of an autonomous ship (Karlis, 2018). Experts have proposed some definitions, typically ranking the level of a ship's autonomy against the amount of human involvement in its operation, such as that provided by Lloyd's register (ShipRight Design and Construction, Additional Design Procedures, Design Code for Unmanned Marine Systems, n.d.).

Prevalence. It is difficult to determine the exact prevalence of drones in seaports and shipping. However, some specific examples exist. For example, Antwerp's Port uses aerial drones for infrastructure inspection, surveillance, berth management, and waste detection (Port of Antwerp deploys autonomous drones for safety enforcement, 2021). The Port of Rotterdam has used drones for shipboard delivery of small cargo. It is even considering drones for offshore shipboard delivery (Drone deliveries take off at the Port of Rotterdam, 2020). The Port of Hamburg uses an Unmanned Surface Vehicle (USV) to monitor water depth autonomously (Howard, 2020). These examples show that autonomous vehicles

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

are employed productively in European ports. However, as these applications are new, the industry will undergo some flux. The industry is recognizing and acting on the potential of drones in shipping.

Green Initiative

The International Maritime Organisation agreed to reduce greenhouse gas emissions by 50% until 2050. To achieve these goals, a consortium of banks has formulated the Poseidon Principles. Signatories integrate climate considerations into their lending decisions. Companies like Patagonia demonstrate significant interest in minimizing their environmental footprint. Cargo owners will show interest in such data and vet ships for their environmental impact.

Pervasive data collection enables every actor to understand their environmental impact. Shippers may select ships on their fuel consumption. Shipowners may identify ineffective ships and replace them. Shipping routes can be optimized for fuel consumption. The data is gathered and combined from various sources. Fuel consumption can be estimated from refueling bills (each gallon of diesel burns to about 22.2 lb. CO2). The data of AIS shows the routes ships have taken. Correlating the data can show routes that require less fuel. The data that we can gather in eShip's system enables us to understand the environmental impact of dry bulk shipping.

Data Analytics

The main issue in analyzing the data is to define the correct model and method that allows predicting the behavior of a system. Researchers are studying methods for extracting information from data sets that are too large to be analyzed with traditional methods. Data mining and machine learning have a significant overlap in methods.

Machine learning aims to predict behaviors, whereas data mining aims to discover properties in data sets. Data that is well structured or minor in the sense that today's computers can process it can use traditional methods. With an SSD of 2 TB capacity on the market, the user should ascertain that the data sets to analyze exceeds 10s of TB before considering big-data methods.

Small data sets have the inherent danger of over-fitting and false discovery rates; the conclusions fit the existing data perfectly but cannot predict anything reliably. Bohannon, Koch, Homm, and Driehaus (2015) demonstrated this phenomenon with their "Chocolate Hoax." By measuring 18 variables in 15 subjects, they established a statistically significant effect of chocolate consumption on weight loss.

Machine Learning

The main methods to consider are statistical methods like linear regression, logistic regression, decision trees, support vector machines, clustering, and machine learning. Dimensionality reduction must be studied for the data sets to make the analysis more tractable. With small data sets, supervised learning is most successful. A group of experts will label the input data to outputs on which the developers train a computer algorithm.

Unsupervised methods leave it to the algorithm to identify structures in the data. These methods allow analysts to discover hidden patterns in the data. They can also learn the features of a particular data set that can be detected in other data items. Reinforcement learning refers to a computer interacting with an environment and receiving feedback in rewards/punishments to direct learning. The algorithm

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

will develop strategies to maximize the rewards/minimize the punishments. Beşikçi, Arslan, Turan, and Ölçer (2016) use a neural network to optimize the choice of ships, routes, and the speed of ships to minimize fuel consumption.

Deep learning using neural networks is the most popular method today. Neural networks aim to simulate how brains are working. A neural network has three layers: an input layer, a recognition layer, and an output layer. Neural networks are trained on sets of input and outputs. The training tunes parameters in the recognition layer that enables the network to produce correct outputs on inputs it was not trained on. Depending on the task, a neural network needs millions of inputs labeled with their outputs. For tasks like image recognition, such neural networks are successful. Language processing is still a significant challenge. Language comprehension is an unachievable goal.

Smart Algorithms Matching Cargo and Ships

Researchers have studied the problem of matching cargo to vessels for multiple decades. They developed the earliest algorithms to maximize the cargo moved on the smallest ships and miles traveled.

Later algorithms generalized the problem of matching containers to ships, optimizing how to load and unload the ship, minimize port time, or minimize shipping cost. In the last decade, work started on matching dry cargo to vessels. Peng, Shan, Guan, and Yu (2016) describe one method for finding a price equilibrium in matching ships to dry cargo. All these matching and optimization are computationally intractable. Thus, fast algorithms will compute suboptimal results. Classes of algorithms used in this context are genetic algorithms, ant-colony optimizations, and recently Bayesian methods. Liang (2019) uses a deep-learning approach to identify transportation patterns and optimize them.

Bilgen and Ozkarahan (2007) use a mixed-integer linear programming model to loading ships optimally using different bulk goods. This allows for blending grains at the harbor and potentially reducing the cost for the customers. Jing, Marlow, and Hui (2008) propose a generalized autoregressive conditional heteroskedasticity (GARCH) model to understand the volatility of prices in the dry bulk shipping market. Similar work was later done by Dai, Hu, and Zhang (2015). The volatility is used to guide investments into ships. Kalouptsidi (2014) analyses the impact of build time on prices in bulk shipping. If ships are in high demand, shipping rates increase. Building more ships to meet the demand may result in decreased rates. Second-hand ships help to reduce prices but may increase the volatility of the investment. Scarsi (2007) analyses market data to demonstrate that shipowners follow their intuition and undervalue market trends. Lun and Quaddus (2009) take a data-driven approach to analyze the shipping market.

Routing and Scheduling

Ronen surveys methods for cargo ship routing and scheduling (Ronen, 1983, 1993). Early methods were optimization methods, for example, the one by Perakis and Bremer (1992). The dissertation of Wu (2020) provides a more recent survey of methods, also considering machine learning-based methods. Norstad, Fagerholt, and Laporte (2011) also consider the vessel's speed as a parameter of a shipping route. This provides the potential to reduce fuel costs and improve the environmental footprint. An energy-efficient fleet has become essential for financing the purchase of new ships. Signatories of the Poseidon principles consider the footprint.

Kang, Zhang, Guo, and Ma (2012) provide an optimization method for ship routing and deployment modes to reduce the operating cost of the shipowner. Guo et al. use data-driven optimization methods

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to assign ships to berths, thus optimizing the time and use of facilities at ports (Guo, Guan & Song, 2012, 2013)

Gao, Sun, Zhao, and Dong (2021) schedules ships for unloading at steel plants. Schim van der Loeff, Godar, and Prakasha (2018) use the AIS to track ships and estimate emissions per vessel. Overall, scheduling methods get deployed in routing and scheduling ships. They enable us to estimate arrival time, reduce shipping and operating costs, and track emissions. Consequently, these methods can provide ways for a greener shipping industry. Mainly, AIS is often transmitted frequently and reliably via satellite. Tracking the position of ships in real-time can be achieved easily by using the services mentioned above.

Processing Documents

Many processes in the shipping industry are still using telefax and written documents. Especially documents of the past are not in digital form. These documents still contain valuable data. Digital documents are often written in prose. While they are comprehensible to humans, processing by computers is challenging. Note that standardized documents are much easier to process for machines but often harder for humans to comprehend without appropriate rendering.

Structured Text Processing

Structured text processing is based on using rules to extract information from specifically formatted text. Text can be processed precisely if it follows the formatting rules. Other texts cannot be understood at all. Structured text processing works well for interpreting forms. The bill of lading is one example for which structured text processing is often sufficient.

Natural Language Processing

Many competitors promise to process e-Mail messages and extract data from those messages. If the messages are not constrained in their structure, the system needs to comprehend the language in the e-Mail.

Some approaches are based on rules, where the text is scanned for specific phrases. Absent these phrases, text cannot be understood precisely. Rule-based methods have similar limitations to structured text processing.

Statistical natural language processing (NLP) combines computer algorithms with machine learning to automatically extract, classify, and label text elements and then assign a statistical likelihood to each meaning of those elements. Deep learning models and learning techniques based on convolutional neural networks and recurrent neural networks enable NLP systems that learn as they work and extract more accurate meaning ever from vast volumes of raw, unstructured, and unlabeled data sets.

SOLUTIONS AND RECOMMENDATIONS

We identify two problem domains that demand specific solutions. The market is heterogeneous, and it is unlikely that market participants will process their data to meet modern standards. If at all, historical data may be digitized and archived. Big data methods allow unlocking these historical data archives by mining them and extracting structured data sets.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Future datasets can be maintained in a standardized format. Such data may be easier to process. However, standards have not been agreed on. eShip is designed to connect to existing ERP systems. The exact mechanism to extract relevant information proposed to access historical data can unify data stored in heterogeneous stores controlled by different entities. Make eShip a responsive platform that provides all the services through desktop, tablet, and mobile phone (Figure 3)

Figure 3. Chartering is just a click away (©2021, eShip project, used with permission)

Unlocking Historical Data

The move to a digital platform creates the problem of converting historical data. This historical data can come from many diverse sources and is in paper form or unstructured electronic documents. Digitizing and converting all this data is a daunting task. Big Data technology allows moving the conversion of the digital assets into the processing and querying of the data. The first step of this proposal is to scan/digitize the documents and store them in a database. The authors suggest that the data is stored unstructured. It should be keyed to unique identifiers. Each data item shall carry a description of its origin and some potential metadata. For example, a bill of lading is often a semi-structured document. A scan of that

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

document can be annotated with the source of the document, keywords "bill of lading," the language it is written in, and the issuing party.

Optical character recognition (OCR) allows for the extraction of the text of these scanned documents. Typewritten and printed text is easy to recognize since the shape of the letters is similar. Handwritten documents occur frequently but are much harder to process. Many documents are created by people involved at the harbor, or printed documents get annotated in handwriting. Natural language processing algorithms can then extract the relevant information from these documents.

FUTURE RESEARCH DIRECTIONS

The bulk shipping industry poses many specific problems that the computing community starts to address right now. The widespread problem is to make historical data, which is often held on paper, available for digital processing. Another problem is forward-looking creating standards to accelerate and automate decision-making in chartering. Both problems need to be addressed individually.

The second area of research addresses process automation. A major problem of charterers and shipowners concerns the scheduling of activities in harbors. Predictive methods for estimating the time and resources needed for these processes reduces cost and increases the utility of the harbor and ships.

CONCLUSION

The most significant risk in choosing a technology is that it does not deliver the value it promises. The authors suggest establishing to define performance metrics to monitor the value processing the data generates. Digitization and automation promise cost savings and faster processes. The standardization of documents and processes will amplify these benefits.

Digitization and standardization are easy to implement for data that is created in the future. Past data is not standardized nor digital. Natural language processing and big data technology enable digital access to historical data. The data is mapped to the digital standards automatically.

At the same time, the field of bulk shipping provides rich research opportunities. Many questions have not been answered yet. We are just starting to become aware of the industry's problems. The maritime industry will benefit from implementing innovative solutions. The authors further claim that those wrong decisions will have a severe impact on the finances of the shipping companies and the environment (Casaca & Loja, 2020). Unlike fully digitized container shipping, bulk shipping, which represents 2/3 of the global maritime transport volume, is in dire need of digital transformation as it still relies on non-structured information exchange. However, bulk shipping is now beginning to adopt a smarter approach to how it best goes about its transformation towards a global value chain. In this chapter, the authors analyzed eShip, an online platform that applies state-of-the-art information and communication technology tools to the traditional chartering business model, generating significant value added to all market participants in three key aspects.

The results point that by providing a digital platform that may be used by all parties involved, either through a web interface or through EDI/PI interfaces, we will be able to maximize interoperability between all the participants, reuse information and reduce management costs and accelerate the digital transition in what is a complex and traditionally conservative sector.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The most important is recognizing the potential of upcoming developments and how data analysis will change operations and create value. The integrated platform with a new vision and collaboration with other parties in the ecosystem will enable stakeholders to assure their positions in the challenging future of the industry that makes its road to become smart. Smart ships and smart ports will exchange data for speeding up operations. With the addition of digital information, data analytics and machine learning uncover better ways to handle cargo, decrease cost, and increase efficiency.

ACKNOWLEDGMENT

The eShip project is supported by the EEA Grants.

REFERENCES

Aiello, G., Giallanza, A., & Mascarella, G. (2020). Towards Shipping 4.0. A preliminary gap analysis. *Procedia Manufacturing*, *42*, 24–29. doi:10.1016/j.promfg.2020.02.019

Autonomous ship project, key facts about YARA Birkeland. (2017). Retrieved from https://www.kongsberg.com/maritime/support/themes/autonomous-ship-project-key-facts-about-yara-birkeland/

Benson, C., D. S., P. & Colling, A. (2018, June). *A Quantitative Analysis of Possible Futures of Autonomous Transport*. Academic Press.

Beşikçi, E. B., Arslan, O., Turan, O., & Ölçer, A. I. (2016, February). An artificial neural network-based decision support system for energy-efficient ship operations. *Computers & Operations Research*, *66*, 393–401. doi:10.1016/j.cor.2015.04.004

Bilgen, B., & Ozkarahan, I. (2007, June). A mixed-integer linear programming model for bulk grain blending and shipping. *International Journal of Production Economics*, *107*(2), 555–571. doi:10.1016/j.ijpe.2006.11.008

Bohannon, J., Koch, D., Homm, P., & Driehaus, A. (2015, December). Chocolate with high cocoa content as a weight-loss accelerator. *International Archives of Medicine*, *8*(55), 1–8.

Cacho, J. L., Marques, L., & Nascimento, Á. (2020). Customer-Oriented Global Supply Chains: Port Logistics in the Era of Globalization and Digitization. In V. Chkoniya, A. O. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 82–103). IGI Global. doi:10.4018/978-1-7998-3115-0.ch005

Casaca, A. C. P., & Loja, M. A. R. (2020). 2019 World of Shipping Portugal. An International Research Conference on Maritime Affairs editorial "Leading the shipping industry into the future." *Journal of Shipping & Trade*, *5*(1), 1–4. doi:10.118641072-020-00082-y

Chen, C.-H., Khoo, L. P., Chong, Y. T., & Yin, X. F. (2014, May). Knowledge discovery using genetic algorithm for maritime situational awareness. *Expert Systems with Applications*, *41*(6), 2742–2753. doi:10.1016/j.eswa.2013.09.042

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chkoniya, V., Madsen, A. O., & Bukhrashvili, P. (2020). *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior*. IGI Global. doi:10.4018/978-1-7998-3115-0

Cross, J., Meadow, G., & Com, I., & Jostaphot. (2017, September). Autonomous ships 101. *Journal of Ocean Technology*, *12*, 23–27.

Dai, L., Hu, H., & Zhang, D. (2015, October). An empirical analysis of freight rate and vessel price volatility transmission in global dry bulk shipping market [English Edition]. *Journal of Traffic and Transportation Engineering*, *2*(5), 353–361. doi:10.1016/j.jtte.2015.08.007

Delenclos, F.-X., Rasmussen, A., & Riedl, J. (2018). *To Get Smart, Ports Go Digital*. Retrieved from https://www.bcg.com/publications/2018/to-get-smart-ports-go-digital

Diwan, M. A. (2016). Internet of Things in Logistics: Towards Autonomous Logistics & Smart Logistic Entities. The International Maritime Transport & Logistics Conference (MARLOG5).

Drone deliveries take off at the port of Rotterdam. (2020, May). Retrieved from https://www.drycargomag.com/drone-deliveries-take-off-at-the-port-of-rotterdam

Ducruet, C., & Zaidi, F. (2012). Maritime constellations: A complex network approach to shipping and ports. *Maritime Policy & Management*, *39*(2), 151–168. doi:10.1080/03088839.2011.650718

Gao, Z., Sun, D., Zhao, R., & Dong, Y. (2021, January). Ship-unloading scheduling optimization for a steel plant. *Information Sciences*, *544*, 214–226. doi:10.1016/j.ins.2020.07.029

Gass, S. I., & Assad, A. A. (2005). *An Annotated Timeline of Operations Research*. Kluwer.

Graziano, M. D., Renga, A., & Moccia, A. (2019). Integration of automatic identification system (AIS) data and single-channel synthetic aperture radar (SAR) images by SAR-based ship velocity estimation for maritime situational awareness. *Remote Sensing*, *11*(19), 16. doi:10.3390/rs11192196

Guo, C., Guan, Z., & Song, Y. (2012). Research on dynamic berth assignment of bulk cargo port based on ant colony algorithm. In Z. Zhang, R. Zhang, & J. Zhang (Eds.), *LISS* (pp. 235–240). Springer. doi:10.1007/978-3-642-32054-5_35

Guo, C., Guan, Z., & Song, Y. (2013). Research on bulk-cargo-port berth assignment based on priority of resource allocation. *Journal of Industrial Engineering and Management*, *6*(1), 276–288. doi:10.3926/jiem.673

Haralambides, H. E. (2019). Gigantism in container shipping, ports and global logistics: A time-lapse into the future. *Maritime Economics & Logistics*, *21*(1), 1–60. doi:10.105741278-018-00116-0

Howard, M. (2020, March). Water drone christened in the port of Hamburg. *Maritime Logistics Professional*.

Jacks, D. S., & Pendakur, K. (2010). Global Trade and The Maritime Transport Revolution. *The Review of Economics and Statistics*, *92*(4), 745–755. doi:10.1162/REST_a_00026

Jing, L., Marlow, P. B., & Hui, W. (2008). An analysis of freight rate volatility in dry bulk shipping markets. *Maritime Policy & Management*, *35*(3), 237–251. doi:10.1080/03088830802079987

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kalouptsidi, M. (2014, February). Time to build and fluctuations in bulk shipping. *The American Economic Review*, *104*(2), 564–608. doi:10.1257/aer.104.2.564

Kang, K., Zhang, W.-c., Guo, L.-y., & Ma, T. (2012, September). Research on ship routing and deployment mode for a bulk. In *2012 international conference on management science & engineering 19th annual conference proceedings.* Dallas, TX: IEEE. 10.1109/ICMSE.2012.6414421

Karlis, T. (2018, January). Maritime law issues related to the operation of unmanned autonomous cargo ships. *WMU Journal of Maritime Affairs*, *17*(1), 1–10. doi:10.100713437-018-0135-6

Kavallieratos, G., Diamantopoulou, V., & Katsikas, S. K. (2020). Shipping 4.0: Security Requirements for the Cyber-Enabled Ship. *IEEE Transactions on Industrial Informatics*, *16*(10), 6617–6625. doi:10.1109/TII.2020.2976840

Lambrou, M., Watanabe, D., & Iida, J. (2019, November). Shipping digitalization management: Conceptualization, typology, and antecedents. *Journal of Shipping and Trade*, *4*(11), 11. Advance online publication. doi:10.118641072-019-0052-7

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, *6*(4), 239–242. doi:10.100712599-014-0334-4

Levinson, M. (2020). *Outside The Box*. Kirkus Reviews.

Liang, Z. (2019, September). Design of automatic matching system for ocean-going cargo in international logistics. *Journal of Coastal Research, 93*(SI), 1105–1110. doi:10.2112/SI93-160.1

Lind, M., Michaelides, M., Ward, R., & Watson, R. T. (2020). *Maritime Informatics*. Springer International Publishing. Retrieved from https://books.google.pt/books?id=Ed6PzQEACAAJ

Lun, Y. V., & Quaddus, M. A. (2009). An empirical model of the bulk shipping market. *International Journal of Shipping and Transport Logistics*, *1*(1), 37. Advance online publication. doi:10.1504/IJSTL.2009.021975

Martimo, P. (2017). *Disruptive Innovation and Maritime Sector - Discovering smart-shipping's potential to disrupt shipping.* Report Turku School of Economics.

Mileski, J., Clott, C., Galvao, C. B., & Laverne, T. (2020). Technical analysis: the psychology of the market of dry bulk freight rates. *Journal of Shipping & Trade, 5*(1). doi:10.118641072-020-00079-7

Munin. (2016). http://www.unmanned-ship.org/munin/

Norstad, I., Fagerholt, K., & Laporte, G. (2011, August). Tramp ship routing and scheduling with speed optimization. *Transportation Research Part C, Emerging Technologies*, *19*(5), 853–865. doi:10.1016/j.trc.2010.05.001

Peng, Z., Shan, W., Guan, F., & Yu, B. (2016). Stable vessel-cargo matching in dry bulk shipping market with price game mechanism. *Transportation Research Part E, Logistics and Transportation Review*, *95*, 76–94. doi:10.1016/j.tre.2016.08.007

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Perakis, A. N., & Bremer, W. M. (1992). An operational tanker scheduling optimization system: Background, current practice and model formulation. *Maritime Policy & Management*, *19*(3), 177–187. doi:10.1080/751248659

Perunovic, Z., & Vidic, J.-P. (2011). Innovation in the Maritime Industry. *Proceedings of the 22nd POMS Annual Conferences*.

Port of Antwerp deploys autonomous drones for safety enforcement. (2021, February). *Dry Cargo International*. Retrieved from https://www.drycargomag.com/port-of-antwerp-deploys-autonomous-drones-for-safety-enforcement

Poulis, K., Galanakis, G. C., Triantafillou, G. T., & Poulis, E. (2020). Value migration: Digitalization of shipping as a mechanism of industry dethronement. *Journal of Shipping and Trade*, *5*(1), 1–18. doi:10.118641072-020-00064-0

Ronen, D. (1983, February). Cargo ships routing and scheduling: Survey of models and problems. *European Journal of Operational Research*, *12*(2), 119–126. doi:10.1016/0377-2217(83)90215-1

Ronen, D. (1993, December). Ship scheduling: The last decade. *Ship scheduling: The last decade, 71*(3), 325-333. doi:10.1016/0377-2217(93)90343-L

Scarsi, R. (2007). The bulk shipping business: Market cycles and shipowners' biases. *Maritime Policy & Management*, *34*(6), 577–590. doi:10.1080/03088830701695305

Schim van der Loeff, W., Godar, J., & Prakasha, V. (2018, December). A spatially explicit data-driven approach to calculating commodity-specific shipping emissions per vessel. *Journal of Cleaner Production*, *205*, 895–908. doi:10.1016/j.jclepro.2018.09.053

Shipwright design and construction, additional design procedures, design code for unmanned marine systems. (n.d.). *Lloyd's Register*. Retrieved from www.lr.org

SmartPort. (2019). Smart ships and the changing maritime ecosystem (T. van Dijk, H. van Dorsser, R. van den Berg, H. Moonen, & R. Negenborn, Eds.). Academic Press.

Stein, M. (2018, June). Integrating Unmanned Vehicles in Port Security Operations: An Introductory Analysis and First Applicable Frameworks. *Ocean Yearbook Online*, *32*(1), 556–583. doi:10.1163/22116001-03201022

Stein, M. (2020, April). *Unmanned maritime infrastructure inspection-a mixed-method risk management approach from German port facilities*. Academic Press.

Tat, C. H., Lim, K., Lund, S., Deggim, H., Koyama, T., Foo, D., . . . de Souza, C. (2020, October). *Future of shipping: Digitalisation webinar.* Retrieved from https://fos-digitalisation2020.sg/

Ward, R., & Bjørn-Andersen, N. (2021). The Origins of Maritime Informatics. In Maritime Informatics. Springer. doi:10.1007/978-3-030-50892-0_1

Wu. (2020). *Bulk ship routing and scheduling under uncertainty* (Doctoral dissertation). Hong Kong Polytechnic University, China. Retrieved from https://theses.lib.polyu.edu.hk/handle/200/10660

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Bulk Shipping: The transportation of goods (either liquid or granular) in large quantity, usually not packed but loaded directly into a vessel. Such goods (bulk cargo) are grains, petroleum products, iron ore, and more.

Chartering: An activity within the shipping industry where a shipowner agrees to rent out his ship to a cargo owner to move cargoes from one point to another.

EEA Grants: Through the European Economic Area (EEA) Agreement, Iceland, Liechtenstein, and Norway are partners in the internal market with the Member States of the European Union. To promote a continuous and balanced strengthening of economic and trade relations, the parties to the EEA Agreement have established a multi-annual Financial Mechanism known as the EEA Grants. The EEA Grants established the objective of reducing social and economic disparities in Europe by strengthening bilateral relations with the Beneficiary States.

Freight Rate: The price requested for the transport of cargo from one place to another. The freight rate depends on the destination, the accountable size of the shipment, and the type of the sheep.

SaaS (Software as a Service): The way of delivering applications over the Internet—as a service. Instead of installing and maintaining software, we access it via the Internet, freeing ourselves from complex software and hardware management.

Shipping Industry: The industry is concerned with transporting freight, especially by ship. A wide variety of commercial transport methods can be found in the shipping industry, from bulk transport of commodities in railcars to highly specialized "intermodal" container shipping. Four significant transport modes exist in this industry: marine, air, rail, and freight (trucking).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 19
Question Answering Chatbot Using Memory Networks

Riddhi Patel
Charotar University of Science and Technology, India

Parth Patel
Charotar University of Science and Technology, India

Chintan M. Bhatt
Charotar University of Science and Technology, India

Mrugendrasinh L. Rahevar
https://orcid.org/0000-0002-0551-9229
Charotar University of Science and Technology, India

ABSTRACT

Manipulation of a large amount of data becomes a very tedious task. Hence, the authors took the approach of memory networks for the implementation of the chatbot. Traditionally, the LSTM model was used to implement chatbots and QA systems. But the LSTM failed to store relevant information when given a longer information set. On the contrary, the memory networks have an additional memory component with it. This can help in storing long information for further use which is greatly advantageous for the QA and chatbot systems as compared to LSTM. The authors trained and tested their model over Facebook's bAbi dataset which consists of several tasks and has questions regarding each task to retrieve the accuracy of the model. On the pedestal of that dataset, they have presented the accuracy for every task in their study with memory networks.

INTRODUCTION

The digital age has bombarded mankind with a large number of devices. The devices stand to retrieve information based on the data they are suspected to. Also, give the users the response they need in the large retrieval-based world of a humongous amount of data that resides in the world. The digital devices

DOI: 10.4018/978-1-7998-6985-6.ch019

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

rely on information retrieval systems such as a chatbot system. The question answering chatbot system strives to retrieve information from a large repository of data with extreme precision and lesser redundancy. This retrieval system is implemented with the foundation of Artificial Intelligence (AI), (Ainouz, S. A. Ben Ahmed, Mohammed, 2020). Also, a chatbot is more like a human and requires intelligence. Artificial Intelligence techniques are essential in its implementation, (M. Lewkowitz, 2014, Feb 12). One of the techniques to be considered as a part of AI is Machine learning (ML). ML in layman terms can be defined as the ability of a machine to learn on its own from the data it is provided and create a prediction or a decision based on the algorithm that is fed into the machine. The chatbot system also requires techniques to mimic a human brain to generate an accurate response, (Bing Liu1, G. T., Hakkani-Tur, P. S. Heck). That is where Deep Learning comes into the picture showing the neural network similar to nerves in the brain of a human.

A chatbot system can be categorized as a retrieval based system and a rule based system, (Jincy Susan Thomas, Seena Thomas, 2018). The proposed system is a chatbot that uses Memory Networks based approach which makes it better than rule-based systems. This system is a Retrieval based system as the bot is trained on a set of questions and their possible outcomes. Memory networks as the name suggests possess an external memory unlike other Neural networks. Having a memory module becomes extremely necessary as the questions sometimes can tend to be so long that a common neural network would fail to backpropagate entirely. Thus, the external memory module can come into play whenever needed in order to backpropagate and process an entire question, (Sainbayar Sukhbaatar, Arthur Szlam Jason Weston Rob Fergus). The instruction set given to the bot makes it possible to get the answer from the dataset it is trained on inorder to get the most relevant answer and output the same. A good and efficiently pre processed dataset can enable the chatbot to produce new answers. Facebook engineers combined a dataset named bAbi inorder to be used as a task response system. The bAbi set consists of 20 tasks that have variable answers, (Jason Weston, Bordes, S.C., Alexander M. Rush, Bart van Merrienboer, Armand Joulin Tomas Mikolov). We have used this dataset to train our chatbot.

The structure of the chatbot that the authors propose has been illustrated in figure 1. The figure 1 shows that whenever an user asks a question, it does the analysis of the question. Further, it retrieves the necessary document that might have an answer for the question for e.g. “where” questions will have answers in “places” documents. Then it retrieves the answer and analyzes it for it’s correctness and finally displays it to the user.

Thus this chatbot has been trained on the machine learning model of memory networks. As, training them end to end requires very less supervision during training which makes it useful in regular scenarios.

BACKGROUND

Evolution of the Chatbot

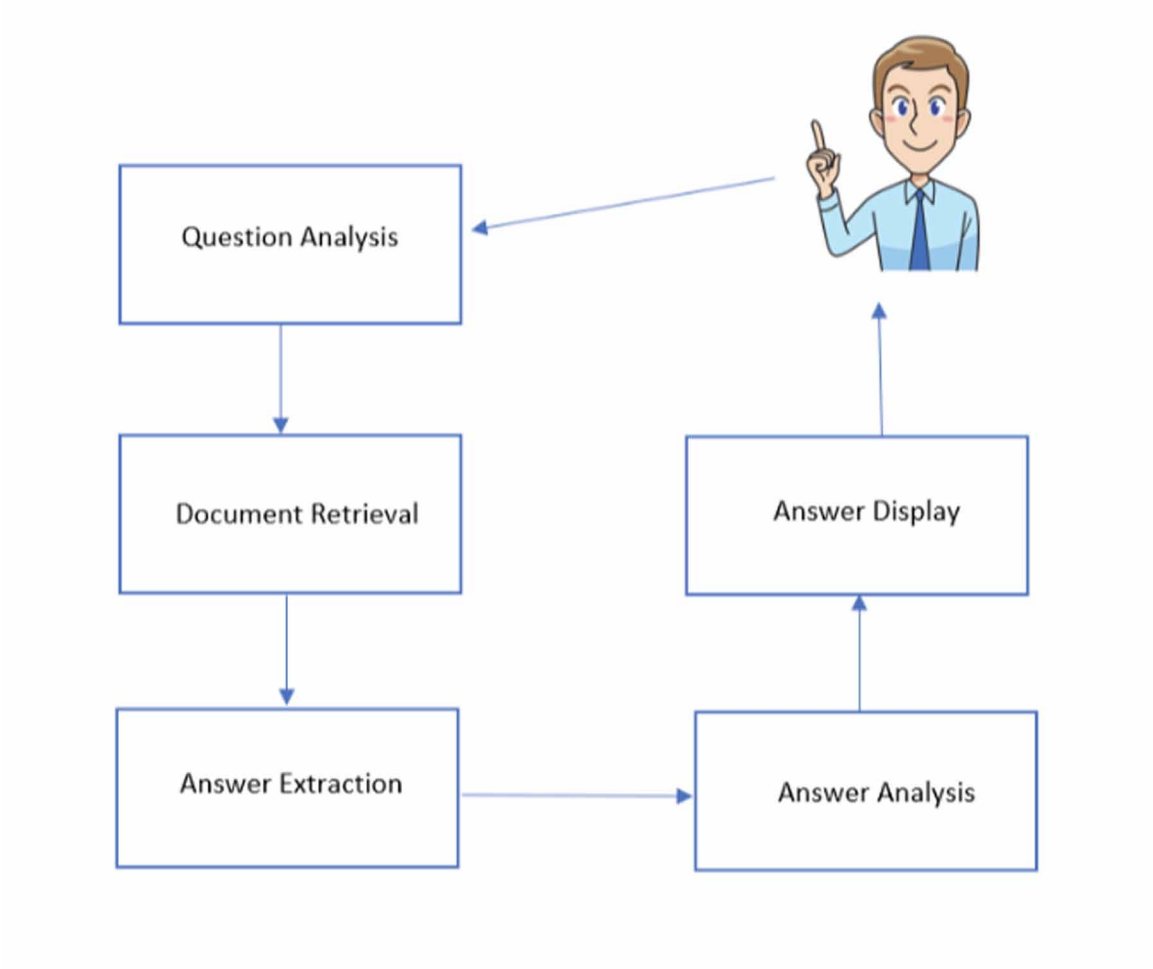
In 1994, Mauldin coined the term ‘ChatterBot’ to define the systems aiming to pass the Turing Test. The first ever chatbot to be made in the computer science field is “Eliza”. Eliza was developed by Joseph Weizenbaum. Eliza’s model captured the inputs and tried to match keywords obtained by rephrasing the input, with predefined responses. The below response can be considered as an example,

utterance = # I want #

Answer = “Why do you want ___?”

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Structure of Chatbot



Here, the "#" is the wildcard character. ELIZA responds with the output string associated with this particular rule out of a set of rules. There are some rules written to rephrase the user utterance before matching. In addition, it also had some responses which were not of any context, possibly used when the input did not match any of the rules. As an example,

utterance = My life is great!

answer = Please go on.

Even though Eliza was created to demonstrate how superficial the conversation between a human and a computer is, ELIZA became widely popular because people thought they were talking to a real person, (M. Lewkowitz, 2014, Feb 12). It was designed to imitate a therapist and many people reviewed that they actually felt relief by talking with Eliza.

Eliza's model works on a rule-based system, which restricts it from having meaningful conversations with humans. Also, as illustrated above, it had generic responses to utterances with which none of the rules matched. Hence, it rather aimed at deceiving the humans into thinking that they were talking to an actual human and not a computer.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

A British programmer named Rollo Carpenter, designed Jabberwacky in 1988. Unlike Eliza, Jabberwacky was aimed to simulate entertaining conversations with humans. It uses an AI technique called "contextual pattern matching". It was one of the first chatbots which harnessed AI to learn meaningful human conversations. Jabberwacky won the Loebner prize, an annual competition to test chatbots, in 2005 and 2006, (Y.Vilner). Jabberwacky was able to learn new facts and concepts just by having constant interactions i.e. it had the ability to learn continuously from the conversations.

From the time Jabberwacky has been online, it has enhanced itself by learning new languages. This has been possible because it relies completely on context and feedback, unlike Eliza. It is purely based on heuristics-based technology, as said by its creators. Infact, the older version of Jabberwacky has now involved in Cleverbot.

A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) is a popular free chatbot developed in 1995 by Dr. Richard Wallace. It is inspired by ELIZA and won the Loebner prize in January 2000, (R.Raine, 2009). The first implementation was known as Program A, which was written in a language based on set theory and mathematical knowledge, known as SETL. A.L.I.C.E was written in AIML, an XML dialect for creating chatbots. AIML was also implemented by Dr. Wallace. When the language migrated to Java, the number of contributors increased and their joint effort launched Program B in 1999. With this Program, A.L.I.C.E. won the Loebner prize in 2000. When Jacob Bikker created the first C and C++ based implementation of AIML, Program C was launched based on it. Finally, Program D led the transition of A.L.I.C.E from pre-Java 2 to add many features such as Swing and Collections. However, A.L.I.C.E failed the Turing Test. It reacts to human inputs and tries to respond as naturally as possible. The XML schema helps simplify the conversational rules. Natural language was the most signif- icant feature of Alice, (alicebot.org).

The bot SmarterChild was launched in 2001. as a service which could answer fact-based questions like "What is the population of Indonesia?" or more personalized questions like "What movies are playing near me tonight?", (R.Raine, 2009). Until 2008, when it was taken down, this chatbot was used by thousands of people daily.

In 2006, IBM Watson developed a chatbot specially to compete with the champions of the game of Jeopardy! It won the competition in 2011, which is when it started offering services to build chatbots for various domains that are capable of processing huge amounts of data.

In 2010, Siri was launched as an intelligent virtual assistant. Later, assistants such as Google Now (2012) by Google, Cortana (2015) by Microsoft and Alexa (2015) by Amazon, were launched inspired by Siri. These assistants are capable of making searches based on Image, easy mobile device access and answer web-based questions. The assistants being commercial, there is no public access to its implementation details.

Conventional Techniques for Question Answering

Prior to the progress of the neural networks, problems of NLP (Natural Language Processing) were given an answer based on traditional approaches such as Named Entity Recognition, Part of Speech Tagging, Stemming Lemmatization, Information Retrieval, Information Extraction. Conventionally rule based approach was used to match patterns of the questions to the answers. Thus providing an output to a comprehension type question by matching rules. The rules are related to being a WH question such as (Who, What, Where, When etc.). This makes it possible to match the question to be of a place type, person type etc. following the rules based on that.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The Question Answering system can be taken into consideration as:

1. Determination of question type
2. Determining answer type based on this
3. NLP tasks are performed
4. System shrinks the corpus to get the precise answer

Comparative Analysis of Question Answering Systems

The computer program pertaining to Question answering system builds it's answer by querying a structured database consisting of information related to websites or any organization and relates the questions to the documents already in natural lan- guage, (Sanjay Dwivedi, Vaishali Singh, 2013). The comparisons for the approaches have been summarized in tables 1, 2 and 3 respectively. Table 1 shows different statistical techniques for question classification, along with the major QA systems that use the systems from their respective websites and their performance. Table 2 shows different techniques that can be used to classify the answers along with major QA systems using it with their performance. Table 3 depicts two approaches for pattern matching over various parameters such as answer extraction, answer representation, etc. for the approaches of surface pattern based and template based.

Table 1. Statistical model for question classification

Statistical Techniques	Major QA system(s)	Performance
Maximum entropy model	IBM's statistical QA	Question classification error rate reduces
Support vector machine	System by Moschitti, Quateroni et al., Zhang et al.	SVM has shown quite good performance and accuracy for question classification and preferred in QA community
Modified Bayesian Classifier	Wei et al., MKQA	Method has better accuracy than base bayesin method

Machine Learning

Humans tend to learn from the environment they are exposed to. As we all know a child learns from the environment it grows up in, similarly machines that are computers tend to learn by the datasets they are exposed to. Machine learning has evolved a lot in the past since the days of its development and the machine learning algorithms have grown exponentially, (Pat Langley).

Learning without any explicit knowledge or input for computers is only made possi- ble by the field of AI known as Machine Learning. It revolutionizes decision making as well as predictions for things done by the computer. Machine Learning can be considered as a child who learns from the environment he is grown up in. It takes the best possible approach to a problem and gives a metric to each of the models that is developed. Thus, Giving the computer an ability like humans.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Statistical model for answer classification

Statistical Techniques	Major QA system(s)	Performance
Okapi Similarity Measurement	IBM's statistical QA	Average level of mean reciprocal ratio for factoid questions in restricted domain
Sentence Similarity Model	System by Cai et al.	Accuracy is significantly increased and above average precision in achieved
N-gram mining	System by Soricut et al.	Satisfactory performance even for the non-factoid questions
SVM	System by Suzuki et al.	SVM outperforms than other models in answer selection phase too

METHODOLOGIES

LSTM (Long Short-Term Memory Networks)

LSTM (Long Short-Term Memory) has the ability to handle long sequences, which is very suitable for a QA system are specialized RNNs that can remember necessary data and forget the irrelevant bits, (Rohit, G. & Dharamshi, Ekta Subramanyam, Natarajan, 2019). The problem with recurrent neural networks (RNNs) was Vanishing Gradients, which limited it to looking back only a few steps in arbitrarily long sequences of information. LSTM mitigates this problem of RNN. Therefore, LSTM is an improvised version of RNN, since it can more efficiently remember longer sequences. It does this with the help of the feature to "remember" the relevant and "forget" the irrelevant parts. Here, the neural network does not have any place to store all the information from the text, making it the biggest problem of LSTM, (Alex Sherstins, 2020).

The advantages of LSTM networks are important, Harman et al argued for the use of supervised learning in contrast to classical NLP methods for answering questions, (Hermann, K.M., Kocisky, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., Blunsom, R, 2015). They describe a model that focuses on the elements of the document, which they believe will help answer the question. It can encode statements in a dataset story. Of these, when questions are placed, it answers, (Xingjian Shi Zhourong

Table 3. Comparison of two pattern matching approaches

	Surface Pattern Based	Template Based
Basic Mechanism	Either human crafted or automatically learned patterns through examples	A template is preformatted framework for questions which have entity slots to be dynamically filled by parameters
Answer Extraction	Answer sentence is extracted using statistical techniques or data mining measures	Uses structured query to extract answers from the database
Answer Representation	Not necessarily formatted answers	Focus on generating formatted answers
Pattern Learning	Semi-automatic	Manually but automatic for semantic web
Most compatible application area	Small and medium size websites	Semantic web

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chen Hao Wang Dit-Yan Yeung). This paper also addresses the lack of good quality datasets to answer the question and the need to artificially generate data to use supervised practice for question answering.

Memory Networks

The biggest challenge that is faced with the machine learning models is to have an easy way to read and write the part of a long component of a memory and to make it in a combination with the inference, (J. Weston, S. Chopra, and A. Bordes, 2015).

Memory Networks basically consists of an array of objects M indexed by M_i. There can be various 4 kinds of components which are known to be learnable.

- I: - Input feature map
  - Using standard processing techniques this map is mainly used to convert input into an internal feature representation.
- G: - Generalization
  - When given new input it converts the old memories.
  - $m_i = G(m_i, I(x), m)$[10]
  - Simplest for is stored in I(x) in $m_{H(x)} = I(x)$, (J. Weston, S. Chopra, and A. Bordes, 2015).
- O: - Output Feature
  - Based on the input x and memory m it produces a new output feature.
  - Output Feature o=O(I(x), m)
- R: - Response
  - Conversion to desired format of the output

Assuming an input of X one can consider the flow of the system as follows:

1. X The internal attribute representation is converted to I (X).
2. Memory mi updated, and new input is given: mi=G (mi, I(X), m) ∀ i, (G. Rohit, E. Dharamshi, Subramanyam, 2019).
3. Output properties o calculated, new input and memory: o=O(I(X), m), (G. Rohit, E. Dharamshi, Subramanyam, 2019).
4. Output properties o decoded to give the final response: r =R (o).

IMPLEMENTATION

Memory Networks

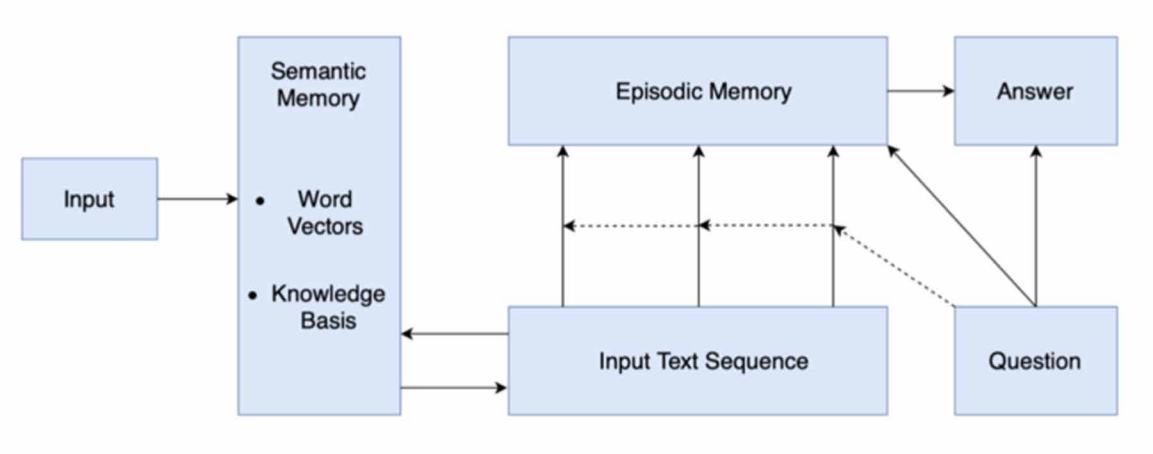
A GRU cell is similar to an LSTM cell, just the difference is that it is more efficient computationally since it has only two gates (update gate and reset gate) and it does not use a memory unit. The update and erase of data is controlled by the two gates. The hidden state in a GRU is calculated by

$h_t = GRU(x_t, h_{t-1})$

Given a work in the previous timestep vector (h_{t-1}), compute the current timestep vector (h_t). The advantage of GRU over LSTM is that it can keep information history without removing the information

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Structure of the memory network



which is irrelevant to the prediction, (K. Ankit, O. Irsoy, J. Su, J. Bradbury, R. English, B. Pierce, P. Ondruska, I. Gulrajani and R. Socher, 2016).

The update gate is a single layer neural network. Now, sum up the matrix multiplications and add a biased term. And then the sigmoid squashes it to a list of values between 0 and 1, the output vector. Do this twice with different sets of weights, then use a rest gate that will learn to ignore the past timesteps when necessary. For example, if the next sentence has nothing to do with those whose came before it. The update gate is similar in that it can learn to ignore the current timestep entirely. Maybe the current sentence has nothing to do with the answer. Whereas, previous ones did.

The question module processes the query word by word and gives output of the vector using the same GRU which had the input module and has the same weight, (C.Xiong,S.Merityand, R.Socher, 2016).

The motivation for this in the chapter came from the hippocampus function in our brain. A hippocampus getting triggered by sigh or sound is able to retrieve the temporal states.

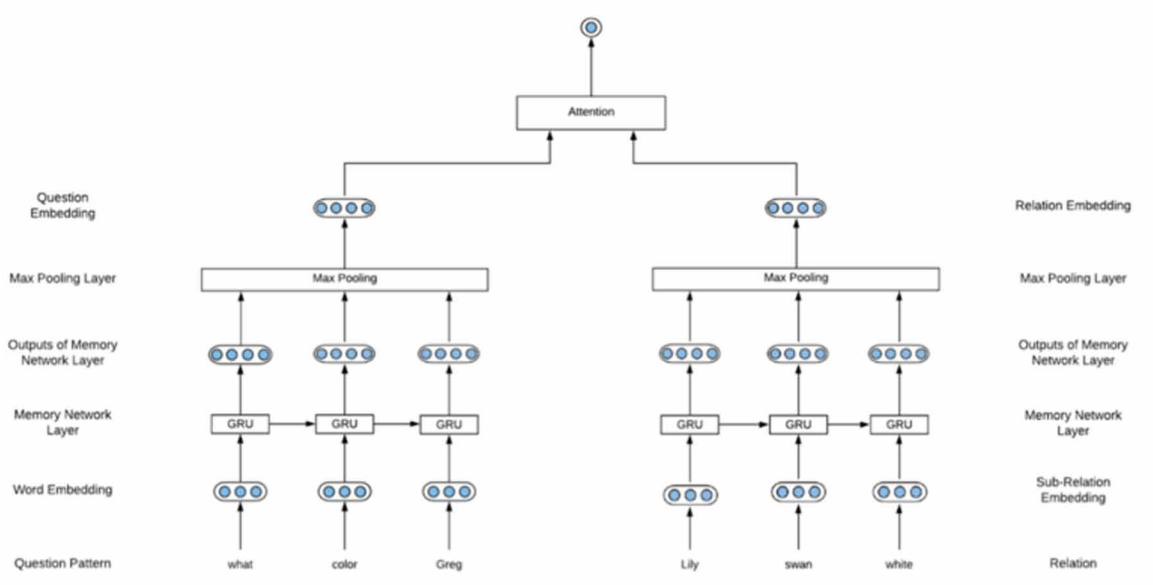
As shown in figure 2, the episodic memory module is made up of two GRUs. It receives the fact and the question vectors that are extracted from the input. Episodes are generated by the GRU on the inner side. It takes into account the output of the attention function based on the current fact when updating its inner state. Each fact is given a score between 0 and 1 by the attention function. And so, the facts with low scores are ignored. At the end of each pass, the outer GRU is given input of episodes which are output of the inner GRU. The purpose of having two layers is for our model to understand after one pass that something else can be important too. Thus, multiple passes can help to gather more information that is correct.

Architecture

The figure 3 illustrates the architectural flow of the question answering system. Here, the question pattern is the question input which is considered to be a sequence of words. These sets of words are then embedded and the sequence of word vec- tors are encoded using GRU. The Max pooling layer provides an abstracted form thus reducing the number of extracted features as well as preventing overfitting. This is concatenated to question embedding thus producing the complete question representation.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Illustration of Architecture



bAbi Tasks Dataset

Facebook prepared a dataset in order to perform various question answering tasks. One such dataset is used in our system. The bAbi tasks dataset comprises 20 differ- ent question answering tasks. The range of the bAbi tasks dataset is from one word answers all the way upto directional tasks. The sample of the dataset is shown in table 4.

RESULTS AND DISCUSSIONS

The Facebook bAbi dataset was used to train and test our model. The tasks listed in table 4 were used by us to benchmark the performance of our model. Figure 4 is a graph representation of the results derived. According to figure 4, the first thing for us is that the accuracy of some tasks is 98% and above,while in others they are less than 30%. Task 3 and 19 turn out to be tasks with the lowest accuracy. Task 3 is the task of three supporting facts, and task 19 is to find the path based on the directions. It is astonishing that Task 2 performed well on two supporting facts, but it was difficult for the model to understand the three supporting facts i.e. finding a way to find a way between two places when the relative positions between each pair is known. The explanation for the performance of the models in this work is that the relative position can be understood in a one-dimensional space, but finding a path that requires two dimensions is difficult.

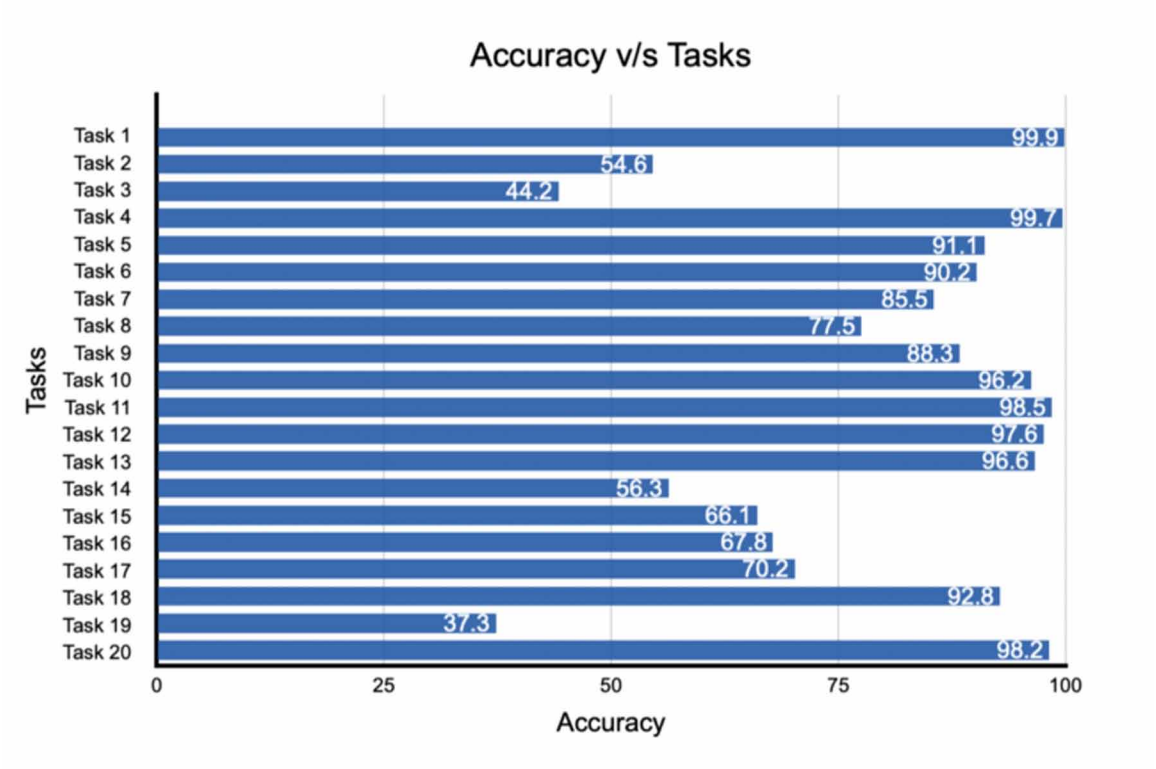
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Sample Dataset

Task 1: Single supporting fact Mary went to the bathroom John moved to the hallway Mary travelled to the office Where is Mary? A: Office	**Task 2: Two supporting facts** John is in the playground John picked up the football Bob went to the kitchen Where is football? A: Playground
Task 3: Three supporting facts John picked up the apple John went to the office John went to the kitchen John dropped the apple Where was the apple before the kitchen? A: Office	**Task 4: Two argument relations** The office is north of the bedroom The bedroom is north of the bathroom The kitchen is west of the garden What is north of the bedroom? A: Office What is the bedroom north of? A: Bathroom
Task 5: Three argument relations Mary gave the cake to Fred Fred gave the cake to Bill Jeff was given the milk by Bill Who gave the cake to Fred? A: Mary Who did Fred give the cake to? A: Bill	**Task 6: Yes/No questions** John moved to the playground Daniel went to the bathroom John went back to the hallway Is John in the playground? A: No Is Daniel in the bathroom? A: Yes
Task 7: Counting Daniel picked up the football Daniel dropped the football Daniel got the milk Daniel took the apple How many objects is Daniel holding? A: Two	**Task 8: Lists/Sets** Daniel picks up the football Daniel drops the newspaper Daniel picks up the milk John took the apple What is Daniel holding? A: Milk, Football
Task 9: Simple negation Sandra travelled to the office Fred is no longer in the office Is Fred in the office? A: No Is Sandra in the office? A: Yes	**Task 10: Indefinite knowledge** John is either in the classroom or the playground Sandra is in the garden Is John in the classroom? A: Maybe Is John in the office? A: No
Task 11: Basic co-reference Daniel was in the kitchen Then he went to the studio Sandra was in the office Where is Daniel? A: Studio	**Task 12: Conjunction** Mary and Jeff went to the kitchen Then Jeff went to the park Where is Mary? A: Kitchen Where is Jeff? A: Park
Task 13: Compound co-reference Daniel and Sandra journeyed to the office Then they went to the garden Sandra and John travelled to the kitchen After that, they moved to the hallway Where is Daniel? A: Garden	**Task 14: Time reasoning** In the afternoon Julie went to the park Yesterday Julie was at school Julie went to the cinema in the evening. Where did Julie go after the park? A: Cinema Where was Julie before the park? A: School
Task 15: Basic deduction Sheep are afraid of wolves Cats are afraid of dogs Mice are afraid of cats Gertrude is a sheep What is Gertrude afraid of? A: Wolves	**Task 16: Basic induction** Lily is a swan Lily is white Bernhard is green Greg is a swan What color is Greg? A: White
Task 17: Positional reasoning The triangle is to the right of the blue square The red square is on top of the blue square The red sphere is to the right of the blue square Is the red sphere to the right of the blue square? A: Yes Is the red square to the left of the triangle? A: Yes	**Task 18: Size reasoning** The football fits in the suitcase. The suitcase fits in the cupboard. The box is smaller than the football Will the box fit in the suitcase? A: Yes Will the cupboard fit in the box? A: No
Task 19: Path finding The kitchen is north of the hallway The bathroom is west of the bedroom The den is east of the hallway The office is south of the bedroom How do you go from the den to the kitchen? A: West, North How do you go from office to bathroom? A: North, West	**Task 20: Agent's Motivations** John is hungry John goes to the kitchen John grabbed the apple there Daniel is hungry Where does Daniel go? A: Kitchen Why did John go to the kitchen? A: Hungry

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Result comparison



CONCLUSION

In this work, the authors have discussed how the addition of an external memory can impact the implementation of a QA chatbot system to a greater extent. The LSTM networks used before were not worthy enough as they used to forget what the question was, making them completely useless for certain smart and security applications. The use of a memory network helps the questions to be processed completely along with certain directional questions. The use of directional questions can be very important by integrating it in smarxt city applications such as tracking the location of a child in a house or a community and also, for security purposes as a question answering chatbot oriented door locks.

REFERENCES

Ainouz, S. A., & Ben Ahmed, M. (2020). *A Smart Chatbot Architecture based on NLP and Machine Learning for Health Care Assistance*. . doi:10.1145/3386723.3387897

Ankit, K., Irsoy, O., Su, J., Bradbury, J., English, R., Pierce, B., Ondruska, P., Gulrajani, I., & Socher, R. (2016). *Ask Me Anything: Dynamic Memory Networks for Natural Language Processing*. ICML.

Bing Liu, Hakkani-Tur, & Heck. (n.d.). *Dialogue Learning with Human Teaching and Feedback in End-to-End Trainable Task-Oriented Dialogue Systems*. Academic Press.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dwivedi, S., & Singh, V. (2013). Modeling Techniques and Applications (CIMTA) Research and reviews in question answering system. *International Conference on Computational Intelligence*. 10.1016/j.protcy.2013.12.378

Hermann, K. M., Kocisky, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., & Blunsom, R. (2015). Teaching machines to read and comprehend. In *28th International Proceedings on Advances in Neural Information Processing Systems* (pp. 1693–1701). MIT Press.

Langley. (n.d.). *Crafting Papers on Machine Learning*. Adaptive Systems Group, Daimler-Chrysler Research and Technology Center.

Lewkowitz, M. (2014, Feb 12). *Bots: The future of human-computer interaction*. Available: https://chatbotsmagazine.com/bots-the-future-of-human-computer-interaction56696f7aff56

Raine, R. (2009). Making A Clever Intelligent Agent: The Theory Behind The Implementation. *IEEE International Conference on Intelligent Computing and Intelligent Systems*. 10.1109/ICICISYS.2009.5358137

Rohit, Dharamshi, & Subramanyam. (2019). Approaches to Question Answering Using LSTM and Memory Networks: SocProS 2017. In *Soft Computing for Problem Solving*. DOI: doi:10.1007/978-981-13-1592-3 15

Rohit, G. (2019). *Dharamshi, Ekta Subramanyam, Natarajan*. Approaches to Question Answering Using LSTM and Memory.

Sherstins, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. *Physica D: Nonlinear Phenomena, 404*.

Sukhbaatar, S., & Arthur, S. J. W. R. F. (n.d.). *End-To-End Memory Networks*. Dept. of Computer Science Courant Institute, New York University.

Thomas & Thomas. (2018). Chatbot Using Gated End-to-End Memory Networks. *International Research Journal of Engineering and Technology, 5.*

Vilner, Y. (n.d.). *Chatbots101:The Evolution of Customer Retention Latest Trend*. Available: https://www.entrepreneur.com/article/293439

Weston, Bordes, Rush, van Merrienboer, Joulin, & Mikolov. (n.d.). Towards Ai-complete Question Answering: A Set Of Prerequisite Toy Tasks. In *Soft Computing for Problem Solving Arch.* Academic Press.

Weston, J., Bordes, A., & Chopra, S. (2015). *Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks*. arXiv preprint, arXiv: 1502.05968

Weston, J., Chopra, S., & Bordes, A. (2015). Memory networks. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.

Xiong, C., Merityand, S., & Socher, R. (2016). *Dynamic Memory Networks for Visual and Textual Question Answering*. ICML.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Accuracy: It is the ratio for the total correct predicted values to the addition of total correct predicted values and total incorrectly predicted values.

Natural Language Processing (NLP): Natural language processing refers to processing of human language for the computer to understand in order to perform/predict according to its needs.

Question Answering System: A question answering system is known for retrieving answers to questions either from structured database, relational database, or unstructured natural language documents.

Recurrent Neural Networks: In a recurrent neural network the output of the previous node is used as an input in the current node. This can help predict the next step for the algorithm.

Vanishing Gradient Problem: With addition of more layers to the neural network the loss function gradient becomes smaller approaching to almost zero which makes it harder to train the network.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 20
Comparison of Parametric and Non-Parametric Methods to Analyse the Data Gathered by a Likert-Type Scale

Tamer Baran
https://orcid.org/0000-0002-8711-6561
Pamukkale University, Turkey

ABSTRACT

The aim of this chapter is to reveal whether the results of the analysis of the data obtained using Likert type scales (LTSs) with parametric and non-parametric methods in different response alternative (DRA) numbers will differ in terms of statistical significance. In this respect, the data were obtained from 271 university students with CETSCALE prepared using LTS in five different response alternatives (DRAs). The data were analysed using the one sample t test and Wilcoxon signed rank test. Significant findings of the study in the analysis of the data obtained using midpoint LTSs and with the normal distribution with both parametric and non-parametric methods couldn't be found. Similarly, the data obtained by four response alternative numbers with the normal distribution were analysed by both methods, and the significant findings were revealed. However, the results of the data obtained by six and eight response alternative numbers with parametric methods were found to be statistically significant while their analysis by non-parametric methods did not reveal significant findings.

INTRODUCTION

The Likert-type scale was put forth by Rensis Likert (1932) in his work titled "A Technique for the Measurement of Attitudes" and is named after him. These types of scales are most commonly used by researchers in many areas ranging from social science to educational science. Likert-type scales are also discussed by applied data scientists in different contexts, besides collecting data. In the context of applied data, whether Likert-type scales are ordinal or interval scales is one of the most controversial

DOI: 10.4018/978-1-7998-6985-6.ch020

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

topics among applied data scientists. This discussion is crucial because the result of the controversy also leads to finding an answer to the question of how to analyze the data obtained from these scales.

Although a number of authors (Aaker et al., 2007; Burns and Bush, 2000; Churchill, 1999; Malhotra, 2004) consider Likert type scales as an example of the interval scale, there are others (for example Jamieson, 2004; Knapp, 1999; Mircioiu and Atkinson, 2017) who claim that Likert-type scales are not interval but ordinal scales. This discussion raises the question of whether the data obtained using Likert-type scales should be tested by parametric or non-parametric methods. Even though there are studies related to this question in different fields such as education (Turan et al., 2015), medicine (Jamieson, 2004; Norman, 2010), and statistics (Sangthong, 2020), the subject has not been discussed enough in the field of social sciences.

Moreover, although Norman (2010) claimed that there is no disadvantageousness in analyzing the data obtained from small samples with Likert-type scales by parametric methods, Wadgave and Kharnier (2016) emphasized that there is a lack of studies to prove this claim. This study was conducted with the motivation to fill this gap in the literature.

The aim of this study is to determine whether the data obtained using Likert-type scales in the different response alternative numbers can be analyzed with parametric and non-parametric methods in terms of statistical significance. For this purpose, the research question of our study is as follows:

Research Question: Would there be a statistical difference in the findings as a result of the analysis of data obtained with Likert-type scales from small samples at different response alternatives with parametric or non-parametric tests?

In order to find an answer to the research question, the chapter is organized into four main sections: In the first section, the literature was reviewed regarding the Likert-type scale. This section has been divided into four sub-sections. In the first sub-section, the emergence and definition of the Likert-type scale are mentioned and information about the use of the Likert-type scale in the literature has been presented. Later, a discussion has been made regarding which type of scales the Likert-type scale conforms to. The last sub-section discusses the different analysis methods suitable for analyzing these scales. In the second main section, the methodology of the study has been explained in detail. In the third main section of this chapter, the findings of the study are presented. In the last section of the chapter, the conclusion, recommendations, and suggestions regarding the topic have been stated.

LITERATURE REVIEW

Likert Type Scales and Using Likert Type Scales in the Literature

Likert type scale, which was first put forward by Rensis Likert in 1932 and named after him, is one of the most frequently used psychometric scales by researchers to evaluate the perception of the participants (Wadgave and Khairnar, 2016). In this type of scale, researchers try to obtain data through numerical values assigned to response alternatives they have created to determine the participants' level of participation to scale items (Chyung et al., 2017). Likert-type scales are one of the most frequently used scales by researchers in many fields, especially due to their benefit in measuring attitudes. Using the Likert scale to measure attitudes and values, researchers can develop their own scales by assigning numbers to levels of participation. The Likert-type scales are used in two ways: (1) for the summarized scale; and (2) for individual items or rating scales from which the aggregated scale is calculated. Likert items are

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

expressions about a specific topic and through these scales, participants are asked to determine their participation levels. The summarized Likert scale is commonly created by developing a series of items about the topic. These items are intended to provide a representative example of all possible views or attitudes on the subject. To calculate the summated scale score, each response type is given a numeric value or weight at specified spaces (Józsa and Morgan, 2017).

Studies about the use of Likert type scales in the literature have focused on some issues including the midpoint (Chyung et al., 2017), response alternative numbers (Baran, 2020), the order of labels (Chyung et al., 2018; Kennedy et al., 2018), the use of positive or negative expressed items in the scales (Chyung et al., 2018), and the presentation of the items to the participant as itemized or scrolled (Chyung et al., 2020).

One of the controversial topics about Likert-type scales is whether the midpoint should be used in scales. While some authors argue that the use of the midpoint in these types of scales would be beneficial, others state that it should not be used. Chyung et al. (2017) emphasize that excluding midpoint in Likert-type scales is not a good idea. In the case of using midpoint, the participants do not feel obliged to channel themselves to one of the extreme values and have the opportunity to present their real ideas on the subject. On the other hand, Likert-type scales with midpoints allow the participants to use the midpoint as a tool to avoid expressing their views on the topic correctly (Chyung et al., 2017). The usage of midpoints in Likert-type scales depends on the researcher's preference (Garland, 1991), the respondent's involvement level to the topic, and the research issue (Johns, 2005; Weems and Onwuegbuzie, 2001).

For researchers who prefer to use the midpoint on Likert-type scales, there is another matter to decide: How should the midpoint be presented? In this regard, researchers have several alternatives such as "*neutral*", "*neither agree nor disagree*", "*undecided*", "*I don't know*", "*it depends*", etc. to present the midpoint. However, researchers should be careful about the use of these options because some words, for example, "*undecided*" may be interrogated whether there is really a middle point between disagreement and agreement, or be seen as a lack of opinion, and for this reason, the participant may be interpreted as confused. Therefore, Chyung et al. (2017) state that terms such as "I don't know" or "depend on" should not be labeled as a midpoint, and should be presented to the participant as an off-scale option if they are really needed. Moreover, the authors express that the most appropriate presentation to be used as the middle point is "*neither agree nor disagree*".

Response alternative numbers of Likert-type scale is one of the most important debated topics in the literature. Although it was first created by Rensis Likert (1932) as 5 response alternatives, over time, the response alternatives have varied from 2 to 101 (Chyung et al., 2017). Researchers must be very careful in determining the number of response alternatives and should use strategies to minimize potential bias. Findings of some studies (Preston and Colman, 2000) show that participants reply easily and quickly when the response alternatives decrease. Similarly, Baran (2020) found that as the number of response alternatives increases, the participants are confused, and consequently the response time and the number of items left blank increase. Moreover, the author mentions that the demographic characteristics of the participants, such as education level, age, etc., are important criteria in determining the number of response alternatives. Also, Baran (2020) emphasizes that the demographics of the participants, such as education level, age, etc., are the criteria to be considered in determining the response number alternatives. However, it must be remembered that a smaller number of response alternatives may result in ignoring options that express the opinions of the participants (Chyung et al., 2017). Therefore, while determining the response alternatives, the researchers should not ignore the characteristics of the respondents, the subject of the study, etc.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

An example of a Likert-type scale with 5 response alternatives is presented in Table 1.

Table 1. A Likert type scale with 5 response alternatives

Item no	Item	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1	A consumer should always buy domestic products.	1	2	3	4	5
2	A real citizen should always buy locally made products.	1	2	3	4	5
3	It is always best to purchase locally products.	1	2	3	4	5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Furthermore, the direction of ranking in Likert-type scales (ascending or descending) is another important issue that should be decided by researchers (Chyung et al., 2018). Table 2 and Table 3 present usage examples of Likert-type scales with response alternatives which are ascending and descending in ranking.

Table 2. Ascending order Likert type scale

Strongly disagree	1	2	3	4	5	Strongly agree

The presentation of the labels in reverse order according to the value given to the label (1- strongly agree, 5- strongly disagree) may cause the participants to be confused.

Table 3. Descending order Likert type scale

Strongly agree	1	2	3	4	5	Strongly disagree

Although a few studies (Christian et al., 2009; Hofmans et al., 2007; Weng and Cheng, 2000) did not find the effect of ranking on participants' response preferences, many studies (Barnette, 2000; Friedman et al., 1993; Hartley and Betts, 2010; Maeda, 2015; Nicholls et al., 2006) found evidence that the direction of response alternatives significantly affects participants' response choice. For this reason, it is

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

important for researchers who use Likert-type scales to pay attention to the response alternatives' orders in terms of the quality of the data to obtain.

Another argument about Likert-type scales is about the positive or negative expression of the items of scales used in the questionnaire. When combined with the numerical values given to the labels briefly mentioned above, the issue becomes even more important. Although negative use of items provides some benefits to researchers, such as helping them reduce acceptance bias and detecting data with acceptance bias, the literature suggests avoiding such usage of items as much as possible. The reason is that using items in this manner may reveal some adverse conditions for researchers. For instance, negatively stated items may appear as a separate factor as a result of the analysis (Greenberger et al, 2003; Ibrahim, 2001; Salazar, 2015). More importantly, careless participants may misunderstand negatively stated items and provide fallacious data (Weem et al., 2003; Woods, 2006). This can undesirably adversely affect data quality in a research application (Chyung et al., 2018). However, in some cases where the attribute to be measured is negative in nature, it may be inevitable to use negative statements in Likert-type scales. In any case, researchers should consider the phenomenon they intend to measure and be careful in using negative statements.

Previous studies (for example, Schriesherim and Hill, 1981; Weem et al., 2006) revealed that the positive or negative use of the items in such scales causes significant changes in the responses of the participants. However, if the researcher still has to use the scale expression negatively, some methods such as presenting the data from negative statements separately (Hartely 2013), creating the negative items clearly (Johnson et al., 2004; Merritt, 2012), warning participants that there are negative statements (Mathews and Shepherd, 2002), and presenting positive and negative statements separately in the scale (Roszkowski and Soven, 2010) can help in gathering quality data.

The Debate as to Whether Likert Type Scales are Ordinal or Interval

One of the most debatable issues of recent periods is whether the Likert-type scale — which is one of the most frequently used scales in marketing research, especially in the measurement of attitudes — is interval scale or ordinal scale (Chyung et al., 2017; Joshi et al., 2015). In this context, it is appropriate to compare interval and ordinal scales first before discussing the Likert-type scales.

Knapp (1999) and Siegel (1957) state that they have two main features that differentiate interval scales from ordinal scales. The first is that the ordinal scales cannot be transformed, while interval scales can be converted linearly. The authors exemplify it with the transformation of C^o temperature into a Fahrenheit unit. Accordingly, heat data obtained from C^o can be transformed into Fahrenheit unit linearly with y= bx + a (F = 1.8 C) o+32). However, this is not the case for ordinal scales. In this context, Goldstein and Hersen (1984) and Tavakoli (2012) emphasized that there is no equal distance between the expressions represented in Likert-type scales, even though mathematically equal intervals are represented. Hart (1996) found that participants don't regard the distance equally in the data obtained using the Likert-type scales. Likewise, according to Kuzon et al. (1996), the statements such as "I strongly disagree, I disagree, I agree and I strongly agree" are examples of an ordinal scale. The levels included in this example can be arranged as strongly disagree < disagree < agree < strongly agree. However, the authors state that the distances between the levels in this arrangement are not equal. Interval scales are clearly defined and consist of evenly spaced levels (Kuzon et al., 1996).

Secondly, there is a specified "0" point on interval scales, while there is no specific "0" point on ordinal scales (Knapp, 1999; Siegel, 1957). For example, the 5-item Likert type scales used in the ques-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

tionnaire forms are generally presented to the participant with the labels "1 = Strongly Disagree", "5 = Strongly Agree" (Jamieson, 2004).

On the Analysis of Data Obtained through Likert Type Scales

The above discussion of whether Likert-type scales are ordinal or interval scales has brought the question of how to analyze the data obtained with this scale. This is the crux of the issue because, as mentioned by Sullivan and Artino (2013), these methods differ in terms of the strength of the analysis. Since the parametric tests can be analyzed using mean and standard deviation, they exhibit stronger results than nonparametric tests. In this context, the usage of one method over the other to analyze the data obtained with Likert-type scales is a problem that should be solved by researchers. On the topic, some authors (for instance Hansen, 2003; Jamieson 2004) argue that the data obtained with these scales should be analyzed with non-parametric methods, while others (Sullivan and Artino, 2013) argue in favor of the parametric methods. On the other hand, there are some authors (Norman, 2010) who claim that this debate is unimportant.

The researchers who claim that Likert-type scales should be tested with nonparametric methods base their argument on the following four reasons (Wadgave and Kharniar, 2016). Firstly, parametric tests require data to be obtained with intermittent or proportional scales and analyze these data using mean and standard deviation. However, some labels used in Likert-type scales, such as occasionally or frequently, do not meet this requirement. For instance, frequently and a half indicates a meaningless finding and is not useful to researchers (Jamieson, 2004). Second, in connection with the above justification, verbal labels assigned to scale items are not suitable for analysis with mean and standard deviation, since they are ordinal scales (Jamieson, 2004). Third, in the use of parametric tests, the sample size and distribution are more important than the level of measurement. Since non-normally distributed data can result in a mean that is not a useful measure in the central trend, therefore, even if Likert type scales are considered interval, data sets obtained with such scales should not be analyzed with parametric tests as they usually have a skewed or polarized distribution (Wadgave and Kharniar, 2016). And finally, the use of labels on a Likert-type scale tends to decrease as they approach the midpoint. In other words, in such scales, participants are more inclined towards the extremes than the midpoints and this causes an effect called the "anchor effect" (Bishop & Herron, 2015). For this reason, using parametric tests are not appropriate in Likert-type scales.

Hansen (2003) and Jamieson (2004) also argue that the Likert-type scales do not meet the above requirements, therefore these types of scales are not interval scales but ordinal ones. On this basis, they have stated that the data obtained using Likert-type scales are not appropriate to be analyzed with parametric tests done with mean and standard deviation. Similarly, Gardner and Martin (2007) expressed that Likert-type scales are not interval scales, but ordinal scales, and that the results of the analysis of the data obtained using these types of scales can only be obtained with non-parametric tests. Boone and Boone (2012) state that in the analysis of data obtained using Likert-type scales, non-parametric methods such as mode, median, and frequency are more appropriate than parametric methods such as mean and standard deviation. Additionally, Bishop and Herron (2015) emphasize that the data should not be obtained using Likert-type scales if the aim is to use parametric methods in the analysis of the research.

On the other hand, some authors argue that there is no harm in using parametric tests in the analysis of data obtained with Likert-type scales. Researchers who defend this opinion present the following reasons. Firstly, the robustness of the parametric tests for the Likert scale is validated not by individual

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

items but by assessing it with a scale that is an aggregated composite score (Carifio and Perla, 2008). Second, in standard t-tests and ANOVA, the data and the central limit theorem work with the normal distribution assumption, showing that the means are normally distributed regardless of the original distribution, unless the sample group is not less than 5 or 10 (Norman, 2010). Third, data simulation studies have revealed the robustness of the Pearson correlation for the sample size between 5 - 60 for the normal and rectangular distribution of the ordinal data (Murray, 2013). Lastly, Monte Carlo studies on data obtained with the Likert scale clearly reveal that the F test is extremely robust even when such data are analyzed at the level of scale items (Wadgave and Kharniar, 2016). Moreover, Norman (2010) claims that the data obtained with Likert-type scales can be analyzed through parametric tests, even if they have small sampling, with non-normally distributed data and unequal variances.

The studies conducted to analyze the data obtained using the Likert-type scales compare with parametric methods or non-parametric methods have revealed different findings. Murray (2013) analyzes the data obtained using 5 category Likert type scales using Pearson, Spearman, and Kendall tau_b correlation methods and asserts that the findings were not statistically different. Norman (2010) transforms the data obtained using Likert-type scales with the 0 - 10 response alternatives into the 1 - 4 form and analyzes using Pearson and Spearman correlation methods. At the end of the study, Norman (2010) finds that the parametric and nonparametric methods didn't reveal different results in terms of statistical significance. Furthermore, Norman (2010) stated that the large sample size, the normal distribution of the data, or the variant equation is not necessary when using parametric tests. Turan et al. (2015) compare the results of the t-test with the Mann-Whitney U test in 27 different samples and found that there were no statistically significant results in 24 samples. In the same study, the authors compare the ANOVA and Kruskal-Wallis methods in 15 different samples and concluded that for 14 samples, there were no statistically significant differences.

Brown (2011), Carifio and Perla (2007), and Carifio and Perla (2008) indicate that Likert-type scales are ordinal scales and that the claim that they should be analyzed with only non-parametric methods is a myth. Furthermore, Carifio and Perla (2008) specify that parametric tests are more appropriate for analyzing the data obtained using Likert-type scales and emphasize that it is best practice to use total scores in analyzing the obtained data with parametric tests. In his study, Vigderhous (1977) analyzes the data obtained using the Likert-type scales, with Pearson as a parametric test, and with Kendall tau_b correlation as a nonparametric test and found that they differed significantly from each other.

Mircioiu and Atkinson (2017) find that in a sample of more than 15 participants, the analysis of data obtained from groups with similar distributions using parametric and non-parametric methods does not vary statistically. Moreover, according to findings, this situation does not change even if the sample size is increased. Therefore, the authors emphasize that the analysis of the data obtained with a Likert-type scale by non-parametric methods may cause some loss of information, and they suggest analyzing the data in question with parametric methods. Sangthong (2020) finds that the analysis of the data obtained by Likert-type scales with 5, 7, and 10 response alternatives in different sample sizes with parametric and non-parametric methods does not cause a difference in terms of statistical significance. Thus, Sangthong (2020) states that it is more plausible to use parametric methods in a situation where there is no difference in terms of significance. However, the findings show that when the assumption of the equivalence of variances is ignored, parametric tests reveal stronger findings than non-parametric tests. Moreover, the findings reveal that the higher the number of response alternatives on these kinds of scales, the stronger the results ensue for both methods.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

METHODOLOGY

The main focus of this study is to reveal whether data collected with Likert-type scales should be analyzed by parametric or non-parametric methods. In order to achieve this aim, the questionnaire form, which was used as a data collection tool in the study, was meticulously prepared by making use of the literature. Although using midpoint in Likert scales is suggested in the literature (Chyung et al., 2017), it was decided to use both types of Likert scales to determine whether the findings differ in terms of midpoint scales and forced scales. As stated in the literature review of the study, there are different alternatives for the researchers in labeling the midpoint. However, since it is emphasized in the literature that the most appropriate label for the midpoint is "neither agree nor disagree", the midpoint in the current study was worded with this label. In this study, the number of response alternatives is one of the most dwelling issues. Although the first usage by Rensis Likert (1932) was with 5 response alternatives, researchers have encountered the use of the Likert scale with many different response alternatives over time. For example, Preston and Colman (2000) expressed that in Likert-type scales, response alternatives were used in a range from 3 to 18, while Chyung et al. (2017) stated that there is a wider usage scale from 2 to 101. Besides, Jamieson (2004) indicated that the most frequently used response alternatives are 5 and 7 in studies that collected data with Likert-type scales. On the use of response alternatives in data collection by Likert-type scales, Jamieson (2004) stated that the most commonly used response alternative number is 5, while Bardakcı (2009) stated that this number is 7 in Turkey. On the other hand, Allen and Seaman (2007) emphasized that the most frequently used response alternatives in researches are 5 and 7. In this context, the data in our study were obtained using Likert type 5; at 5±1 and 7; 7±5 different number alternatives (4, 5, 6, 7, and 8).

To determine the ordering direction of the responses to the items in the questionnaire form and the values given to the labels, the findings of previous studies were considered. Moreover, in order to minimize the potential confusion that participants may experience, the response alternatives were presented horizontally and were created in ascending order. Finally, in previous studies, based on the findings on the effect of the negative presentation of the items on the response of participants, attention was paid to the positive use of expressions in the scale.

In this study, Consumer Ethnocentrism Scale (CETSCALE) which was developed by Shimp and Sharma (1987) was used as a scale. This scale was chosen in the study due to the fact that increasing the response rate inattentively reduces the possibility of response (Galesic and Bosnjak, 2009: 357). The 5-point Likert type CETSCALE, one of the questionnaire forms used in the study, is presented at the end of the chapter in Table 7.

The data of the study were collected by convenience sampling. This method was used because it enables data collection in the easiest, fastest, and most economical way (Hasiloglu et al, 2015). The questionnaire forms with different response alternatives were applied to small samples. There were two reasons for this. First, the purpose of the study was to gather data in small samples. The second was the effort not to act incongruously the independence of observations assumption. In this context, data was collected from 271 university students by face-to-face survey method and was analyzed through the One-Sample t-Test as a parametric method and the Wilcoxon Signed Rank Test as a non-parametric method. The findings are presented in the following section.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Reliabilities of Scales by Alternative Numbers

Number of Response Alternative	Cronbach's α
4	.878
5	,847
6	.907
7	.861
8	.930

Table 5. Results of the Kolmogorov-Smirnov Test for different response alternative numbers

Numbers of Response Alternative	N			K-S Z Test	p
4	59	2,75	0.4541	,636	,813
5	56	3.05	0.5805	,787	,571
6	47	4.13	0,8750	,714	.687
7	63	3.83	0.9779	,648	,795
8	46	5,86	1.3521	.697	,715

Table 6. Comparison of One Sample t Test and Wilcoxon Signed Rank Test Results in terms of Response Alternative Numbers

Number of Response Alternative	Type of Test	N	$\bar{x}$ / $\tilde{x}_{median}$	$Sh_{\bar{x}}$	Test results		
					T / Z values	*ss*	*p*
4	One-Sample *t* Test	59	2,75	,0591	4,280	0.4541	< **,001**
	Wilcoxon Signed Rank Test		3	-	-1,976	-	<,**048**
5	One-Sample *t* Test	56	3.05	,0776	0,693	0.5805	,491
	Wilcoxon Signed Rank Test		3	-	-1,100	-	,271
6	One-Sample *t* Test	47	4.13	,1276	32,365	0,8750	< **,001**
	Wilcoxon Signed Rank Test		4	-	-1,630	-	,103
7	One-Sample *t* Test	63	3.83	,1245	-1,334	0.9779	,187
	Wilcoxon Signed Rank Test		4	-	-1,156	-	,248
8	One-Sample *t* Test	46	5,86	,1994	6,824	1.3521	< **,001**
	Wilcoxon Signed Rank Test		6	-	-0,880	-	,379

FINDINGS

The Cronbach's α coefficients, which show the reliability of the CETSCALE scale used in the study according to the number of different response alternatives, are given in Table 4. As can be seen from

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the Table, all scales are presented with different alternative numbers that are reliable (Nunnally, 1978). Besides, the highest α coefficient belongs to the scale where the 8-alternative number was used.

The Kolmogorov-Smirnov test results for the distribution of the obtained data are shown in Table 5. It can be seen that all of the data obtained using different alternative numbers are normally distributed.

According to different alternative numbers of the study, the findings for One-Sample t and Wilcoxon Rank test results are shown in Table 6.

One-Sample *t*-Test and Wilcoxon Signed Rank Test results show that the data obtained using Likert type scales with midpoint (consisting of 5 and 7 labels) were not statistically significant as well as the test conducted with both methods. Besides, findings also show that the parametric test results of the data obtained using Likert-type scales consisting of 4 response alternatives from forced scales were statistically significant and 99% of nonparametric test results were statistically significant. On the other hand, the data obtained using Likert-type scales consisting of 6 and 8 response alternatives yielded statistically significant results for the parametric tests, whereas it didn't yield statistically significant results for nonparametric tests.

CONCLUSION

Likert-type scales are the most widely used scales to measure consumer attitudes in marketing research. It is one of the controversial issues of recent periods whether the distance between the levels in the Likert type scales are not equal and therefore cannot be transformed and whether such scales will be evaluated as interval scale due to the lack of a certain point "0". This discussion brought the question of whether the data obtained using such scales should be analyzed by parametric or nonparametric methods.

In this context, this study tried to find out whether the results of the data obtained using Likert-type scales by testing with two methods would differ in terms of statistical significance. In this regard, data were gathered from 271 university students in 5 different response alternatives by face-to-face survey method. In the analysis of the data, One-Sample t-Test was used as the parametric method and the Wilcoxon Rank Signed Test was used as the nonparametric type of this method.

As a result of the analyses, it was found that there was no difference in the statistical significance in the case of analyzing by means of the midpoint (consisting of 5 and 7 response alternatives) and by means of the parametric or nonparametric method for the normally distributed data. Similarly, it was found that the findings were not different in terms of statistical significance for the data gathered with the scale of 4-response alternative. Statistically significant findings were obtained in the data normally distributed for both parametric and nonparametric methods. These findings support the findings of Murray (2013), Norman (2010), and Turan et al. (2015). In the data collected with Likert-type scales consisting of 6 and 8 response alternatives, the results of the analysis with the parametric tests were found to be statistically significant, whereas the results of the analysis with nonparametric tests were not statistically significant. These findings support the findings of Vidgerhous (1977).

Our findings are similar to the findings of previous studies for different response alternatives when evaluated on the basis of data obtained from different response alternative numbers. Moreover, the findings of our study clearly show that the decision to analyze the data obtained using the Likert-type scales with parametric methods or non-parametric methods doesn't depend on the number of response alternatives.

However, in this study, it should not be overlooked that all of the data obtained using the different number of response alternatives shows the normal distribution. Furthermore, the equation of variants

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

was not analyzed due to the methods of this study. Therefore, it is useful for the researchers to take into consideration the suggestions of the literature at the decision stage when deciding to use parametric methods or nonparametric methods in their studies.

REFERENCES

Aaker, D. A., Kumar, V., & Day, G. S. (2008). *Marketing research* (9th ed.). John Wiley & Sons.

Aji, B. M., & Larner, A. J. (2017). Screening for dementia: Single yes/no question or Likert scale? *Clinical Medicine*, *17*(1), 93–94. doi:10.7861/clinmedicine.17-1-93 PMID:28148591

Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, *40*(7), 64–65.

Awang, Z., Afthanorhan, A., & Mamat, M. (2016). The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences*, *4*(1), 13.

Baran, T. (2020). Anket Formuyla Veri Toplamada Renk ve Cevap Alternatifi Sayısının Cevaplama Süresi ve İfade Cevaplama Sayısına Etkisi. *Akademisyen Yayınevi*. *Bölüm*, *5*, 97–116.

Bardakcı, A. (2009). Pazarlama Araştırmalarında Kullanılan Tutum Ölçeklerindeki Cevap Alternatifi Sayısına İlişkin Bir Literatür Taraması. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (4), 7–20.

Barnette, J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*(3), 361–370. doi:10.1177/00131640021970592

Batterton, K., & Hale, K. (2017). The Likert Scale What It Is and How to Use It. *Phalanx*, *50*(2), 32–39.

Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. *International Journal of Exercise Science*, *8*(3), 297. PMID:27182418

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, *50*(2), 1–5.

Brown, J. D. (2011). Likert items and scales of measurement. *Statistics*, *15*(1), 10–14.

Burns, A., & Bush, R. (2000). *Marketing Research* (3rd ed.). Prentice Hall.

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x PMID:19120943

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106–116. doi:10.3844/jssp.2007.106.116

Christian, L. M., Parsons, N. L., & Dillman, D. A. (2009). Designing scalar questions for Web surveys. *Sociological Methods & Research*, *37*(3), 393–425. doi:10.1177/0049124108330004

Churchill, G. A., Jr. (1999). Marketing Research Methodological Foundations (7th ed.). The Dryden Press International Edition.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chyung, S. Y., Barkin, J. R., & Shamsy, J. A. (2018). Evidence-Based Survey Design: The Use of Negatively Worded Items in Surveys. *Performance Improvement*, *57*(3), 16–25. doi:10.1002/pfi.21749

Chyung, S. Y., Hutchinson, D., & Shamsy, J. A. (2020). Evidence-Based Survey Design: Ceiling Effects Associated with Response Scales. *Performance Improvement*, *59*(6), 6–13. doi:10.1002/pfi.21920

Chyung, S. Y., Kennedy, M., & Campbell, I. (2018). Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert-Type Response Options. *Performance Improvement*, *57*(9), 9–16. doi:10.1002/pfi.21800

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement*, *56*(10), 15–23. doi:10.1002/pfi.21727

Chyung, S. Y., Swanson, I., Roberts, K., & Hankinson, A. (2018). Evidence-based survey design: The use of continuous rating scales in surveys. *Performance Improvement*, *57*(5), 38–48. doi:10.1002/pfi.21763

De Winter, J. C., & Dodou, D. (2010). Five-point Likert items: T test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, *15*(11), 2.

Friedman, H. H., Herskovitz, P. J., & Pollack, S. (1993). The biasing effects of scale-checking styles on response to a Likert scale. In Proceedings of the American Statistical Association annual conference. *Survey Research Methods*, *2*, 792–795.

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, *87*(3), 564–567. doi:10.1037/0033-2909.87.3.564

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. doi:10.1093/poq/nfp031

Gardner, H. J., & Martin, M. A. (2007). Analyzing ordinal scales in studies of virtual environments: Likert or lump it! *Presence (Cambridge, Mass.)*, *16*(4), 439–446. doi:10.1162/pres.16.4.439

Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing Bulletin*, *2*(1), 66–70.

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, *35*(6), 1241–1254. doi:10.1016/S0191-8869(02)00331-8

Hansen, J. P. (2003). CAN'T MISS—Conquer Any Number Task by Making Important Statistics Simple. Part 1. Types of Variables, Mean, Median, Variance, and Standard Deviation. *Journal for Healthcare Quality*, *25*(4), 19–24. doi:10.1111/j.1945-1474.2003.tb01070.x PMID:14606209

Hart, M. C. (1996). Improving the Dissemination Of SERVQUAL By Using Magnitude Scaling. In *Total Quality Management Action* (pp. 267–271). Springer. doi:10.1007/978-94-009-1543-5_42

Hartley, J. (2013). Some thoughts on Likert type scales. *International Journal of Clinical and Health Psychology*, *13*, 83–86.

Hartley, J., & Betts, L. (2010). Four layouts and a finding: The effects of changes in the order of the verbal labels and numerical values on Likert type scales. *International Journal of Social Research Methodology*, *13*(1), 17–27. doi:10.1080/13645570802648077

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Haşıloğlu, S. B., Baran, T., & Aydın, O. (2015). Pazarlama araştırmalarındaki potansiyel problemlere yönelik bir araştırma: Kolayda örnekleme ve sıklık ifadeli ölçek maddeleri. *Pamukkale İşletme ve Bilişim Yönetimi Dergisi*, (1), 19–28.

Hofmans, J., Theuns, P., Baekelandt, S., Mairesse, O., Schillewaert, N., & Cools, W. (2007). Bias and changes in perceived intensity of verbal qualifiers effected by scale orientation. *Survey Research Methods*, *1*(2), 97–108.

Ibrahim, A. M. (2001). Differential responding to positive and negative items: The case of a negative item in a questionnaire for course and faculty evaluation. *Psychological Reports*, *88*(2), 497–500. doi:10.2466/pr0.2001.88.2.497 PMID:11351897

Ivanov, O. A., Ivanova, V. V., & Saltan, A. A. (2018). Likert scale questionnaires as an educational tool in teaching discrete mathematics. *International Journal of Mathematical Education in Science and Technology*, *49*(7), 1110–1118. doi:10.1080/0020739X.2017.1423121

Jacoby, J., & Matell, M. S. (1971). Three-point Likert scales are good enough. *JMR, Journal of Marketing Research*, *8*(4), 495–500. doi:10.1177/002224377100800414

Jamieson, S. (2004). Likert scales: How to (ab) use them. *Medical Education*, *38*(12), 1217–1218. doi:10.1111/j.1365-2929.2004.02012.x PMID:15566531

Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion, and Parties*, *15*(2), 237–264. doi:10.1080/13689880500178849

Johnson, J. M., Bristow, D. N., & Schneider, K. C. (2004). Did you not understand the question or not? An investigation of negatively worded questions in survey research. *Journal of Applied Business Research*, *20*(1), 75–86.

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *Current Journal of Applied Science and Technology*, 396-403.

Józsa, K., & Morgan, G. A. (2017). Reversed items in Likert scales: Filtering out invalid responders. *Journal of Psychological and Educational Research*, *25*(1), 7–25.

Kennedy, M., Campbell, I., & Chyung, Y. (2018). *Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert Response Options*. Academic Press.

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, *39*(2), 121–123. doi:10.1097/00006199-199003000-00019 PMID:2315066

Likert, R. (1932). A technique for the measurement of attitudes. *Archives de Psychologie*.

Maeda, H. (2015). Response option configuration of online administered Likert scales. *International Journal of Social Research Methodology*, *18*(1), 15–26. doi:10.1080/13645579.2014.885159

Malhotra, N. K. (2004). *Marketing Research an Applied Orientation* (Vol. 4). Pearson, Prentice Hall.

Mathews, B. P., & Shepherd, J. L. (2002). Dimensionality of Cook and Wall's (1980) British Organizational Commitment Scale revisited. *Journal of Occupational and Organizational Psychology*, *75*(3), 369–375. doi:10.1348/096317902320369767

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Merritt, S. M. (2012). The two-factor solution to Allen and Meyer's (1990) Affective Commitment Scale: Effects of negatively worded items. *Journal of Business and Psychology*, *27*(4), 421–436. doi:10.100710869-011-9252-3

Mircioiu, C., & Atkinson, J. (2017). A comparison of parametric and non-parametric methods applied to a Likert scale. *Pharmacy (Basel, Switzerland)*, *5*(2), 26. doi:10.3390/pharmacy5020026 PMID:28970438

Murray, J. (2013). Likert data: What to use, parametric or non-parametric? *International Journal of Business and Social Science*, *4*(11).

Nicholls, M. E., Orr, C. A., Okubo, M., & Loftus, A. (2006). Satisfaction guaranteed: The effect of spatial biases on responses to Likert scales. *Psychological Science*, *17*(12), 1027–1028. doi:10.1111/j.1467-9280.2006.01822.x PMID:17201782

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice*, *15*(5), 625–632. doi:10.100710459-010-9222-y PMID:20146096

Nunnally, J. C. (1978). *Psychometric Theory*. Academic Press.

Pell, G. (2005). Use and misuse of Likert scales. *Medical Education*, *39*(9), 970–970. doi:10.1111/j.1365-2929.2005.02237.x PMID:16150039

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. doi:10.1016/S0001-6918(99)00050-5 PMID:10769936

Roszkowski, M., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, *35*(1), 117–134. doi:10.1080/02602930802618344

Salazar, M. S. (2015). The dilemma of combining positive and negative items in scales. *Psicothema*, *27*(2), 192–199. PMID:25927700

Sangthong, M. (2020). The Effect of the Likert Point Scale and Sample Size on the Efficiency of Parametric and Nonparametric Tests. *Thailand Statistician*, *18*(1), 55–64.

Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement*, *41*(4), 1101–1114. doi:10.1177/001316448104100420

Shimp, T., & Sharma, S. (1987). Consumer ethnocentrism: Construction and validation of the CETSCALE. *JMR, Journal of Marketing Research*, *26*(August), 280–289. doi:10.1177/002224378702400304

Siegel, S. (1957). Nonparametric statistics. *The American Statistician*, *11*(3), 13–19.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science. New Series*, *103*(2684), 677–680.

Sullivan, G. M., & Artino, A. R. Jr. (2013). Analyzing and interpreting data from Likert type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. doi:10.4300/JGME-5-4-18 PMID:24454995

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Turan, I., Şimşek, Ü., & Aslan, H. (2015). Eğitim araştırmalarında likert ölçeği ve likert tipi soruların kullanımı ve analizi. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, (30), 186–203.

Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, *47*(1), 65–72.

Vickers, A. J. (2005). Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. *BMC Medical Research Methodology*, *5*(1), 35. doi:10.1186/1471-2288-5-35 PMID:16269081

Vigderhous, G. (1977). The level of measurement and "permissible" statistical analysis in social research. *Pacific Sociological Review*, *20*(1), 61–72. doi:10.2307/1388904

Wadgave, U., & Khairnar, M. R. (2016). Parametric tests for Likert scale: For and against. *Asian Journal of Psychiatry*, *24*, 67–68. doi:10.1016/j.ajp.2016.08.016 PMID:27931911

Weem, G. H., Onwuegbuzie, A. J., & Collins, K. M. T. (2006). The role of reading comprehension in responses to positively and negatively worded items on rating scales. *Evaluation and Research in Education*, *19*(1), 3–20. doi:10.1080/09500790608668322

Weems, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement & Evaluation in Counseling & Development*, *34*(3), 166–176. doi:10.1080/07481756.2002.12069033

Weng, L., & Cheng, C. (2000). Effects of response order on Likert type scales. *Educational and Psychological Measurement*, *60*(6), 908–924. doi:10.1177/00131640021970989

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factory analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 189–194. doi:10.100710862-005-9004-7

Zhang, X., & Savalei, V. (2016). Improving the factor structure of psychological scales: The Expanded format as an alternative to the Likert scale format. *Educational and Psychological Measurement*, *76*(3), 357–386. doi:10.1177/0013164415596421 PMID:27182074

ADDITIONAL READING

Batterton, K., & Hale, K. (2017). The Likert Scale What It Is and How to Use It. *Phalanx*, *50*(2), 32–39.

Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. *International Journal of Exercise Science*, *8*(3), 297. PMID:27182418

Brown, J. D. (2011). Likert items and scales of measurement. *Statistics*, *15*(1), 10–14.

De Winter, J. C., & Dodou, D. (2010). Five-point Likert items: T test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, *15*(11), 2.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jamieson, S. (2004). Likert scales: How to (ab) use them. *Medical Education*, *38*(12), 1217–1218. doi:10.1111/j.1365-2929.2004.02012.x PMID:15566531

Mircioiu, C., & Atkinson, J. (2017). A comparison of parametric and non-parametric methods applied to a Likert scale. *Pharmacy (Basel, Switzerland)*, *5*(2), 26. doi:10.3390/pharmacy5020026 PMID:28970438

Murray, J. (2013). Likert data: What to use, parametric or non-parametric? *International Journal of Business and Social Science*, *4*(11).

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice*, *15*(5), 625–632. doi:10.100710459-010-9222-y PMID:20146096

Pell, G. (2005). Use and misuse of Likert scales. *Medical Education*, *39*(9), 970–970. doi:10.1111/j.1365-2929.2005.02237.x PMID:16150039

Wadgave, U., & Khairnar, M. R. (2016). Parametric tests for Likert scale: For and against. *Asian Journal of Psychiatry*, *24*, 67–68. doi:10.1016/j.ajp.2016.08.016 PMID:27931911

KEY TERMS AND DEFINITIONS

Likert Type Scale: Likert scale is a method which was developed by Rensis Likert and used to measure participants' attitudes towards any issue.

Non-Parametic Test: In statistics, the analysis methods applied using mod, median etc. when the data unmeet certain criteria.

Parametric Test: The analysis methods applied using the mean and standard deviation, in case the data meet certain criteria, in statistics.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX

Table 7. 5-point Likert type CETSCALE, one of the questionnaire forms were used in the study

	Kesinlikle katılmıyorum	**Katılmıyorum**	**Ne katılıyorum ne katılmıyorum**	**Katılıyorum**	**Kesinlikle katılıyorum**
Her zaman en doğrusu Türk ürünlerini satın almaktır.	1	2	3	4	5
Bir Türk yabancı ürün satın almamalıdır, çünkü bu Türk ekonomisine zarar verir.	1	2	3	4	5
Yabancı ürün satın alan Türk tüketiciler diğer Türklerin işsiz kalmasından sorumludur.	1	2	3	4	5
Yabancı ürün satın almak doğru değildir, çünkü bu Türkleri işsiz bırakır.	1	2	3	4	5
Uzun vadede benim için daha maliyetli olsa da Türk mallarını satın almayı tercih ederim.	1	2	3	4	5
Çok gerekli olmadıkça diğer ülkelerden ürün satın alınmamalıdır.	1	2	3	4	5
Başka ülkelerin mallarını satın alarak onları zenginleştirmek yerine Türk mallarını satın almalıyız.	1	2	3	4	5
Gerçek bir Türk her zaman Türk markalı ürünleri satın alır.	1	2	3	4	5
Yabancıların Türkiye'de ürün satmalarına izin verilmemelidir.	1	2	3	4	5
Bir Türk her zaman ithal ürünler yerine Türk ürünlerini tercih etmelidir.	1	2	3	4	5
Benim için Türk yapımı ürünler her zaman önceliklidir.	1	2	3	4	5
Türkiye'ye girişini azaltmak için yabancı ürünlerin vergileri artırılmalıdır.	1	2	3	4	5
Sadece ülkemizde üretilmeyen ürünler yabancı ülkelerden alınmalıdır.	1	2	3	4	5
Yabancı markalı ürün satın almak Türkiye karşıtlığıdır.	1	2	3	4	5
Türk ürünlerini satın alarak Türkiye'nin üretimi artırılmalıdır.	1	2	3	4	5
Sadece Türkiye'de üretilmeyen ürünler ithal edilmelidir.	1	2	3	4	5
Yabancı ürünlerin ithalatı engellenmelidir.	1	2	3	4	5

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 21
Latent Fingerprint Enhancement Based on EDTV Model

Shadi M. S. Hilles
https://orcid.org/0000-0002-2605-9524
Istanbul Okan University, Turkey

Abdilahi Deria Liban
Hargeisa University, Somalia

Abdullah M. M. Altrad
Al-Madinah International University, Malaysia

Yousef A. Baker El-Ebiary
Universiti Sultan Zainal Abidin, Malaysia

Mohanad M. Hilles
University College of Applied Science, Palestine

ABSTRACT

The chapter presents latent fingerprint enhancement technique for enforcement agencies to identify criminals. There are many challenges in the area of latent fingerprinting due to poor-quality images, which consist of unclear ridge structure and overlapping patterns with structure noise. Image enhancement is important to suppress several different noises for improving accuracy of ridge structure. The chapter presents a combination of edge directional total variation model, EDTV, and quality image enhancement with lost minutia re-construction, RMSE, for evaluation and performance in the proposed algorithm. The result shows the average of three different image categories which are extracted from the SD7 dataset, and the assessments are good, bad, and ugly, respectively. The result of RMSE before and after enhancement shows the performance ratio of the proposed method is better for latent fingerprint images compared to bad and ugly images while there is not much difference with performance of bad and ugly.

DOI: 10.4018/978-1-7998-6985-6.ch021

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Fingerprint as one of Biometrics types is widely used in these days. It's one of the most frequently used biometrics to identify individuals and authenticate their identity. It has been used for the recognition for more than 100 years (Galbally et al., 2018). It's commonly categorized into three different categories as presented by (Jiangyang *et al*., 2013) which are rolled, plain and latent fingerprints. The process that obtained in each category is different from the rest. For instance, rolled is achieved by rolling the finger from one side to the other with the purpose of capturing all the ridge particulars. The plain fingerprint is accomplished by pressing the fingertip onto a flat surface. The last category, Latent fingerprints normally lifted from object surfaces that were touched by criminal in the crime scenes.

Biometric System Phases

The literature on biometric phases (Gupta *et al*, 2020) reported that biometric identification system contains two different fundamental phases. The first phase is enrolment while the second is recognition.

In the first phase, a sensor normally gathers the biometric data from which a set of features are extracted and kept in a database along with the individual's identity for instance name, identification number, and birthdates. In the second phase, the identity of the person is either confirmed (verification) or determined (identification). This is achieved by collecting the biometric data, extracting the same features and comparing them to the features stored in the database.

From this comparison, similarity score is produced to make a decision to whether the two sets of features came from the same person or not. The following figure 1 shows an overview of a biometric system phases.

The above figure 1 present an overview of a biometric system phases that consist enrolment and recognition. The process in each phase is clearly defined. Another study documented by (Sabhanayagam *et al*., 2018) demonstrate that all biometric systems comprises of three fundamental elements as per the followings: -

Enrollment: It is the process of gathering biometric samples from an individual. This person is known as the enrollee, and the succeeding generation of his/her template.

Templates: it can be defined as the data representing the enrollee's biometric.

Fingerprint is significant used method for identification systems which is a unique feature such as ridges and valleys where is ridge is a collection of minutiae points included core points as per discussed and mentioned on Menon, et al (2015) and within fingerprint classes as illustrated on figure 2 there are a common three structures which are Arch, Loop and Whorl

Structures Noise in Latent Fingerprint Images

According to (Paulino *et al*., 2013) the main difference between latent fingerprint images and rolled or plain fingerprints is the presence of structured noise in latent fingerprint images. We can categorize the structured noise that exists in latent fingerprints images into six main categories including arch, line, character, speckle, stain and others. The following figure 3 presents the six types of structured noise.

Based on the above figure 3 structured noise can be classified into six main categories as in the elaboration below:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Overview of a biometric system

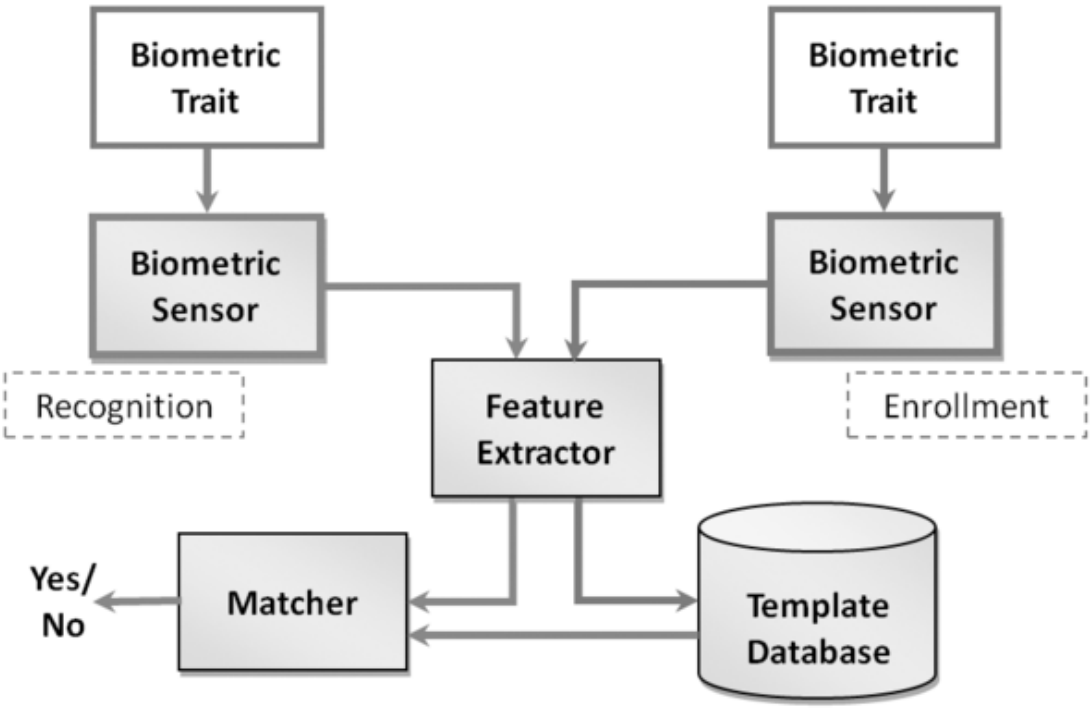
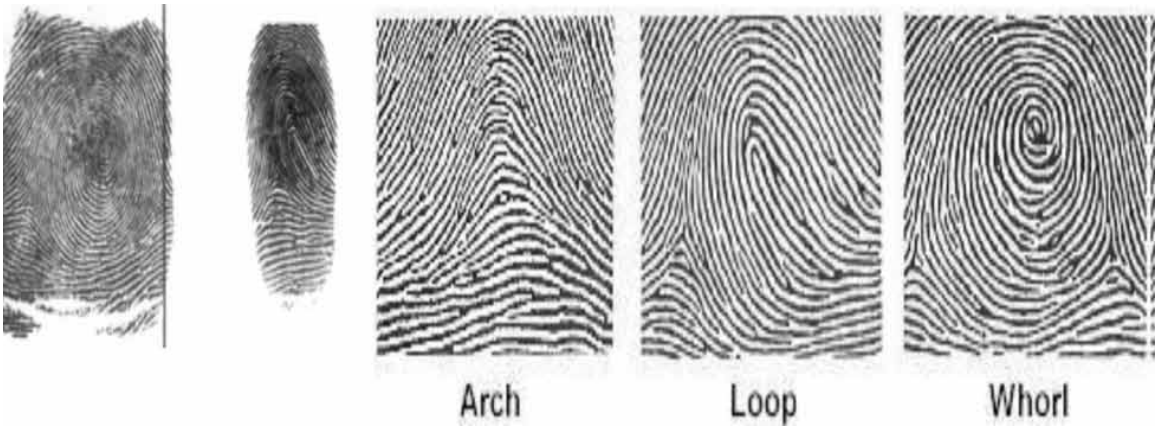
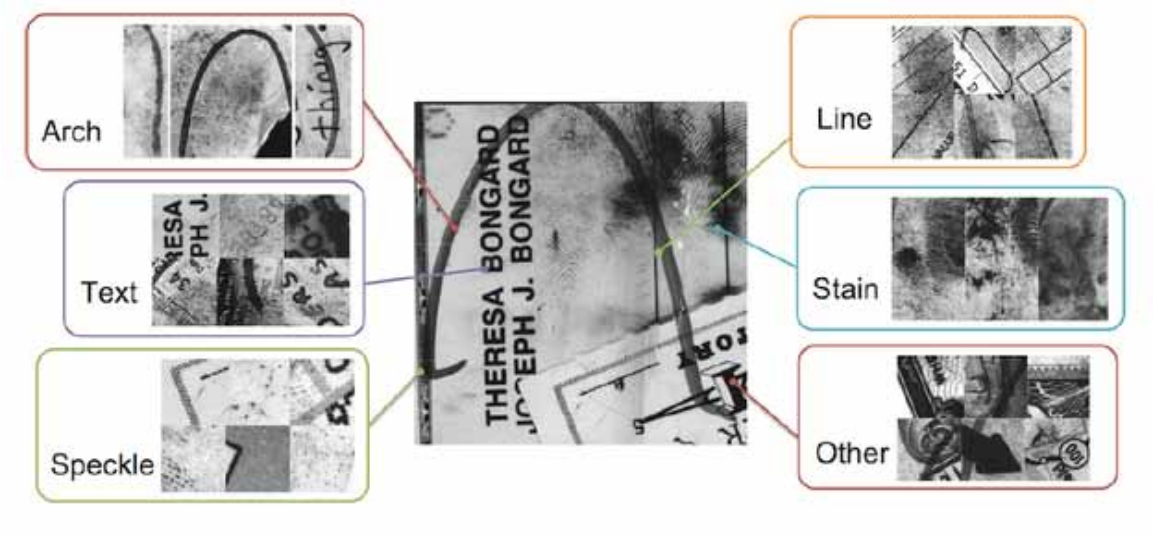


Figure 2. Three Classes of fingerprints



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Six types of structured noise



- Arch: The arch is manually marked by crime-scene investigators to show the possible existence of latent fingerprints in the region encircled by the arch. This category (arch) noise is viewed as the simplest type of structured noise.
- Line: The line noise appears in form of a single line or multiple parallel lines. A single line is often perceived and removed using methods based on the Hough transform (Mukhopadhyay *et al*, 2015). Multiple parallel lines can be confused with fingerprints more easily since they share quite a few common features.
- Character: This is one of the most common types of structured noise in latent fingerprints. Characters may appear in various font types, sizes and brightness that can be either handwritten or typed.
- Stain: This type of latent fingerprint noise is generated when the fingerprint was inadvertently smeared on a wet or dirty surface. Stain noise often appears in spongy shape with inhomogeneous brightness.
- Speckle: Compare to lines and characters noises, speckle noise tends to contain tiny-scale structures, which are either regular such as clusters of small dots or random like ink and dust speckles.
- Others: A latent fingerprint images could comprise other types of structured noise. For example, arrows, signs, etc. Being similar to arch noise and character noise, they usually consist of smooth surfaces with sharp edges.

Usually the line, character and speckle noise groups often come into view when the latent fingerprint is lifted from the surface of a text document like maps, newspapers and checks. For latent fingerprint enhancement, the major challenge lies in how to successfully divide latent fingerprints, the relatively weak signal, from all categories of the noise in the background, which is often the dominant image component. Besides, extra complexity comes when structured noise overlaps with the fingerprint signal.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Latent Fingerprint Enhancement Techniques

The literature on fingerprint enhancement shows variety of approaches has been proposed to solve the issue of quality of fingerprint images. Previous research has shown the application of Gabor filter in fingerprint enhancement as proposed by (Zahedi et al, 2015). In this technique, the local ridge orientation and frequency are firstly approximated in every pixel. After that a Gabor filter is tuned to the local orientation and frequency and applied on the image pixel to suppress the unwanted noise and improve the clarity of ridge structure. The major drawback of this technique is that it needs consistent estimation of local ridge orientation and frequency. This process is challenging for poor quality fingerprint especially latent fingerprint images. Several techniques has been proposed also including methods based on spatial domain, the significant classification of fingerprint enhancement and identification problems such as lack of reliable minutia extraction (Khan et al, 2016) frequency domain, development fingerNET based on convolutional neural network CNN which given significant performance where is CNN extracted fingerprint features and remove structure noise in term of enhance fingerprint by multi task learning (Li et al, 2018)

It is highly important to highlight that the above discussed techniques were mainly proposed for the rolled and plain fingerprints and little attentions have been given to latent fingerprint images. Thus, these mechanisms don't work well on latent fingerprint due to the poor-quality latent fingerprint with smudged, blurred and broken ridge structures. Furthermore, latent fingerprint is usually overlapped with various types of structured noise such as arch, line, character and handwriting. As result, reliable estimation of local ridge orientation and frequency is a difficult task in latent fingerprint. Therefore, the above approaches may not be practical in this situation.

For several years' great effort has been devoted to the study of plain and rolled fingerprint enhancement while there has been little systematic analysis of latent fingerprint image enhancement/pre-processing. Its relatively new area requiring research in order to put on firmer foundations, using extraction of minutia points, fingerprint enhancement consider ridge pattern which is essential pre-processing step thus needs to reduce false positive and negative detection rates (Svoboda, 2017).

Table 1 illustrated the most relevant work of latent fingerprint enhancement, the summaries of using techniques and limitation, and shows the advantages of using Gabor filter and EDTV.

Methodology

Research methodology adopted in designing the improved model for latent fingerprint recognition/matching. This part is organized into seven sections. It begins by highlighting the problem situations and the respective solution concepts and follows by an overview of the research methodology. also discusses the SD27 latent fingerprint dataset used in this study. Design considerations for improving latent fingerprint matching are explained. The metrics for evaluating the performance of the model are given towards the end of the part together with the methodology for results analysis and validation. Finally, the summary of the chapter is concluded.

Table 2 shows the main problem of this chapter and its solution concept based on hybrid technique of latent fingerprint enhancement, and as per discussion in the literature review, image enhancement is required for latent fingerprint images to improve the accuracy of the proceeding steps such as latent fingerprint image segmentation and matching. The main goal of latent fingerprint enhancement is to improve the clarity of the ridge structures while the major challenging problem for latent fingerprint

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Latent Fingerprint Enhancement Techniques

No	Author	Title	Technique and Advantages	Limitation
1	Yoon *et al*, 2010	On latent fingerprint enhancement	Gabor filter method that uses STFT method has been introduced to attain several orientation elements for each block.	Requires region of interest (ROI) singular points and uses fixed ridge frequency
2	Feng *et al*, 2011	Separating overlapped fingerprints	Suggested technique to guess orientation field. The technique applies the previous knowledge of fingerprint ridge structure. "It represented by a dictionary of reference orientation patches and the compatibility constraint between neighboring orientations patches".	Requires region of interest (ROI) and uses fixed ridge frequency
4	(Zhang *et al.* 2013)	Adaptive directional total-variation model for latent fingerprint segmentation	The authors additionally recommended an adaptive directional total variation (ADTV) model. They have combined both local orientation and scale of latent fingerprint enhancement and segmentation.	Relies on orientation field and orientation coherence estimation. One of the drawbacks is that, not an easy task to consistently guess the local parameters such as ridge orientation and scale of adaptive total variation (ADTV) model for poor quality latent fingerprint images. Furthermore, the noise corrupted regions are not restored as well. In addition to that, the extracted fingerprint pattern is usually very weak. Therefore, this will limit the performance of latent fingerprint matching.
5	Qian, P., Li, A., & Liu, M. (2019)	Latent fingerprint enhancement based on DenseUNet	A Gabor filter	Needs consistent estimation of frequency and local ridge orientation. This is normally very challenge for poor quality fingerprint such as latent fingerprint images.
8	Feng et al. 2013	Orientation field estimation for latent fingerprint enhancement	Robust method to estimate orientation field	One of the drawbacks is the fixed ridge frequency.
9	Liu, S., Liu, M., & Yang, Z. (2017)	Sparse coding-based orientation estimation for latent fingerprints	Based on ridge orientation and scale	It is not an easy task to reliably estimate the local parameters like ridge orientation and scale of adaptive total variation model for latent fingerprint images of poor quality. Furthermore, the noise corrupted regions are not restored as well. In addition to that, the extracted fingerprint pattern is usually very weak. Therefore, this will limit the performance of latent fingerprint matching.
10	Cao *et al*, 2014	Segmentation and Enhancement of Latent Fingerprints: A Coarse to Fine Ridge Structure Dictionary	This technique is based on dictionary. It allows reliable estimation of frequency and ridge orientation. It facilitates the automatic segmentation and enhancement of latent fingerprints. Latent fingerprint enhancement mechanism, that effectively integrates the total variation model and the multi-scale patch, based sparse representation for eliminating noises and enlightening the clarity of ridge structure has been recommended.	There are two drawbacks in this approach. Foremost, in the regions of high curvature, the assumption of a single dominant ridge orientation is not valid. Consequently, the Gabor filters with fixed orientation will be likely to destroy the ridge structure and lead to spurious ridge arte facts. Secondly, though the Gabor filtering with correct orientation and frequency parameters can work well to enhance the ridge clarity, it fails to restore the ridge structure destroyed by heavy structured noises that exist in latent fingerprint images. The recommended mechanism can be further enhanced by using of the global ridge structure through global optimization for the low-quality latent fingerprints. "Although reliable estimation of ridge orientation and frequency can improve the performance of latent fingerprint enhancement, there are two inherent limitations in these methods. First, in the regions of high curvature, the assumption of a single dominant ridge orientation is not valid. As a result, the Gabor filters with fixed orientation will be likely to destroy the ridge structure and lead to spurious ridge artefacts. Second, although the Gabor filtering with correct orientation and frequency parameters can work well to enhance the ridge clarity, it fails to restore the ridge structure destroyed by heavy structured noises"
11	T.Revathy, 2014	Automatic Latent Fingerprint Segmentation based on Orientation and Frequency Features	Orientation and frequency features based latent fingerprint segmentation.	The technique can be further improved by including other features like ridge types, curvature minutia types and position of minutia points.
12	Zhang *et al*, 2013	Adaptive directional total-variation model for latent fingerprint segmentation	Proposed adaptive total variation model to remove the structured noises for latent fingerprint segmentation, authors further proposed an adaptive directional total variation (ADTV) model by integrating the local orientation and scale for fingerprint segmentation and enhancement"	One of the drawbacks of these techniques is that, it's not easy task to consistently estimate the local parameters such as ridge orientation and scale of adaptive total variation model for latent fingerprint images of poor quality. Furthermore, the noise corrupted regions are not restored and the extracted fingerprint pattern is usually very weak. Thus, this reduces and limits the performance of latent fingerprint matching.
13	Liu, M., Liu, S., & Yan, W. (2018).	Latent fingerprint segmentation based on ridge density and orientation consistency	Uses ridge orientation tensor and local Fourier analysis.	Does not generate the desired output for the entire fingerprint since some unwanted signals will be extracted for the noise region.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summary of problem situation and solution concepts

No	Problem situation	Solution concept
1	Low quality of latent fingerprint images captured in crime scenes	Apply hybrid technique based on edge adaptive directional total variation model with lost minutia reconstruction technique to enhance the quality of images.

enhancement is to eliminate a variety of types of image noises while consistently restoring the corrupted regions and enhancing the ridge clarity. Thus, this chapter discusses the result of the proposed solution of the latent fingerprint enhancement which is based on edge adaptive directional total variation model with lost minutia reconstruction technique to enhance the quality of latent fingerprint images.

Contains two main research activities which concern on latent fingerprint image de-noising and lost minutia construction.

Latent fingerprint image enhancement receives an input image, applies set of intermediate on the input image and finally outputs the enhanced image.

Figure 4. Latent fingerprint enhancement Research Framework

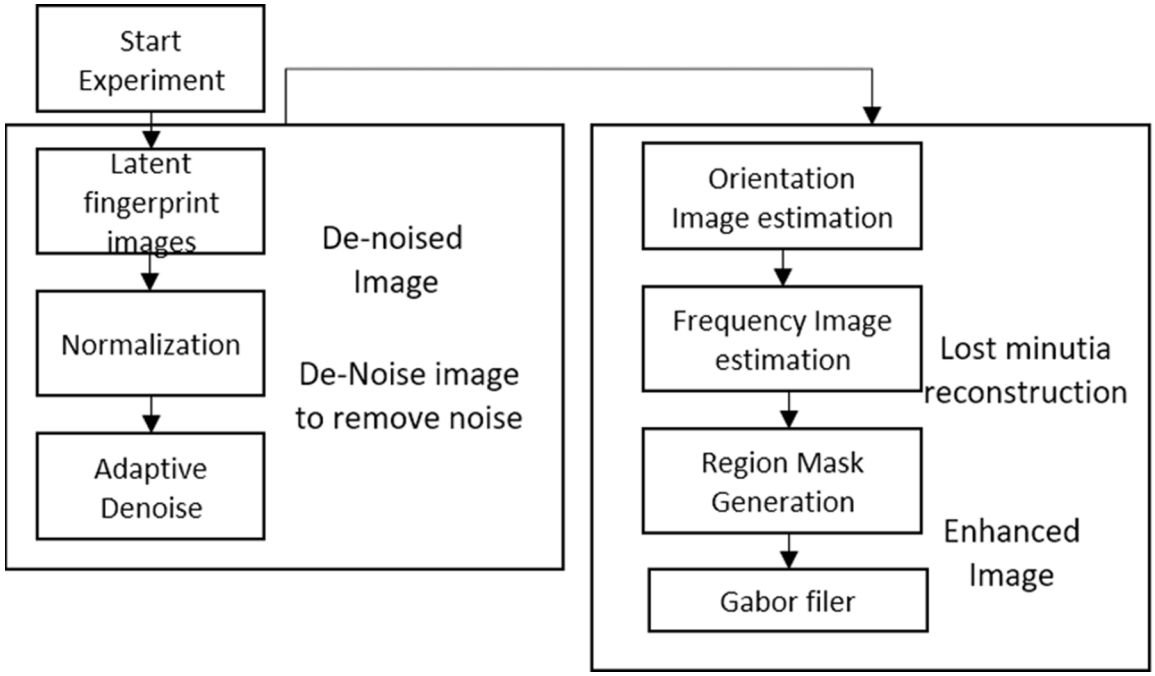


Figure 4 proposed hybrid latent fingerprint enhancement a gray-level latent fingerprint image, where is I, is defined as $\mathrm{N}\times\mathrm{N}$ matrix, where $\mathrm{I}(i,j)$ represents the intensity of a pixel at the ith row jth column. The mean and variance of gray level latent fingerprint image I are defined as

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$M(\mathrm{I}) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\mathrm{I}(i,j)\ and \tag{1}$$

$$VAR(\mathrm{I}) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(\mathrm{I}(i,j) - M(\mathrm{I}))^2 \tag{2}$$

Equation 1 and 2 shows the orientation image, ϑ, is defined as $\mathrm{N}\times\mathrm{N}$ image, where $\vartheta(i,j)$ represents the local ridge orientation at pixel (i,j). Local ridge orientation is usually specified for a block rather than every pixel, an image is divided into set of $\omega\times\omega$ non overlapping blocks and a single local ridge orientation is defined in each block.

A frequency image, $\mathcal{F}$, is defined as $\mathrm{N}\times\mathrm{N}$ image, where $\mathcal{F}(i,j)$ represents the local ridge frequency which is defined as the frequency of the ridge and furrow structures in a local neighborhood along a direction normal to the local ridge orientation.

The region mask $\mathcal{R}$ is defined as $\mathrm{N}\times\mathrm{N}$ image with $\mathcal{R}(i,j)$ indicating the category of the pixel. Pixel could be either:

1. non-ridge-and-furrow (unrecoverable) pixel with value 0 or
2. ridge and furrow recoverable pixel with value 1

Normalization: input latent fingerprint image is normalized so that it has pre-specified mean and variance to reduce the variations in gray level values along ridges and valleys. This is to facilitate the subsequent processing steps. The normalization process specifies mean and variance. Let, $\mathrm{I}(i,j)$ denote the gray-level value at pixel (i,j), M and VAR denote the estimated mean and variance of I respectively, and $G(i,j)$ denote the normalized gray-level value at pixel (i,j). The normalized image is defined as follows where $\mathrm{M0}$ and $VAR0$ are the desired mean and variance values respectively, it's a pixel wise operation and does not change the clarity of ridge and furrow structures as shown on equation 3.

$$G(x,y) = \begin{cases} M0 + \sqrt{\dfrac{VAR0(\mathrm{I}(x,y)-M)^2}{VAR}}, if\ \mathrm{I}(x,y) > M \\ M0 - \sqrt{\dfrac{VAR0(\mathrm{I}(x,y)-M)^2}{VAR}}, otherwise \end{cases} \tag{3}$$

Adaptive de-noise based on Edge Directional Total Variation (EDTV) model. It is used to effectively remove the noise of latent fingerprint images. The model is adaptive to image edge direction and capable of handling images with several dominant directions. To recover an image f(x,y) from a given noisy image i(x, y) Rudin *et al* proposed the Total variation Model which can be summarized in the following mathematical expression:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$E_{TV}\left(f\right)=\lambda\int\Omega\mid\nabla f\mid d\Omega+\frac{1}{2}\int\Omega\left(f-1\right)^{2}d\Omega \tag{4}$$

Bayram and Kamasak further improved the model by proposing directional total variation model for the image demonizing (DTV).

$$E_{DTV}=\lambda DTV_{\alpha,\theta\left(f\right)+\frac{1}{2}}\int\Omega\left(f-1\right)^{2}d\Omega \tag{5}$$

Equation 4 and 5 illustrated the comparison between total variation and directional total variation where is DTV model enhances the diffusion along the direction θ when the dominant direction in an image coincides with the direction θ. It's necessary to make the parameter θ spatially varying throughout the entire image when there are several dominant directions in the image. Thus, Hua Zhang, and Yuanquan Wang proposed Edge Directional Total Variation Model (EDTV) which supports spatially varying θ(x,y) based on the edge direction of the image as follows:

$$\left(\cos(\theta(x,y,\sin\left(\theta\left(x,y\right)\right)\right)=n_{1}\left(x,y\right),n_{2}\left(x,y\right) \tag{6}$$

Equation 6 shows cosine of edge direction, Where n_1 (x,y), n_2 (x,y) is the edge direction. The DTV model in 2 with θ(x,y) in 3 is refered to as EADT Model. In this way the EADTV model is adaptive and enhances the diffusion along the image edge direction. The direction of the edge in latent fingerprint noisy images n_1 (x,y), n_2 (x,y) is very important and should be calculated in advance (n_1 (x,y), n_2 (x,y)) = (-gy, gx) / $g^2_x + g^2_y$. Where (gy, gx) is the gradient factor of g(x,y) and g(x,y) is a smoothed version of I(x,y) using a guising filter.

RESULT AND EXPERIMENT

Presents the result of RMSE of good latent fingerprint images in SD27 Dataset. The average of RMSE for the total 85 good images in SD27 dataset is 0.018373 In addition to that, the result of RMSE of latent fingerprint bad images is slightly lower.

Table 3. RMSE Average of Image categories

Image Category	RMSE Average
Good	0.018373
Bad	0.022287
Ugly	0.023199

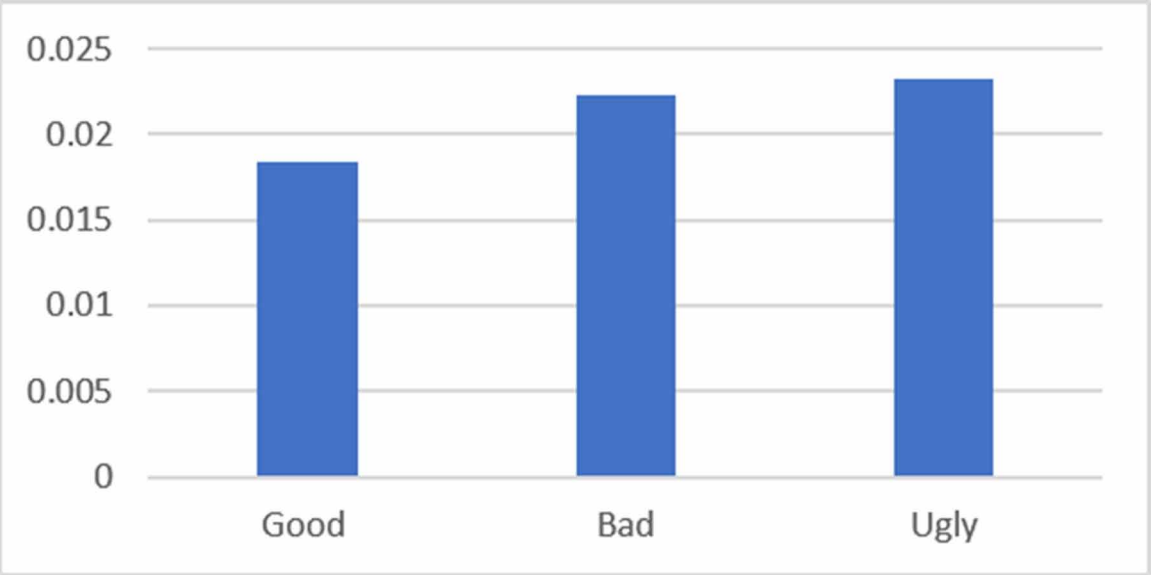
Table 3 shows the average of RMSE of three image categories is compared. The following figure 5 present the comparison of RMSE for SD27 image categories.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. ANOVA Test for RMSE

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.001	2	.001	21.208	.000
Within Groups	.007	252	.000		
Total	.008	254			

Figure 5. RMSE comparisons for SD27 Image Categories



From figure 5, the comparison of RMSE for SD27 image categories has been illustrated. The lesser the number of RMSE the better the performance of the result. Therefore, the proposed solution performs better for good, bad and ugly images respectively.

It's important to highlight as well that ANOVA test is used to observe the differences of performance between the latent fingerprint image groups, good, bad, and ugly images available in SD27 data set.

The table 4 shows the output of the ANOVA of RMSE for three categories analysis and whether we have a statistically significant difference between our means. We can see that the significance level is 0.00 ($p = .00$), which is below 0.05. Therefore, there is a statistically significant difference in the mean length of RMSE for different categories of the images (good, bad and ugly). This is great to know. However, we do not know which of the specific categories differed since the result doesn't indicate which of the three categories differ from one another. Thus, we can find this out in the multiple comparisons table which contains the results of post-hoc tests as in the following table 5.

The above tables presented the comparison of RMSE for different categories of images. The result shows that the proposed technique for the latent fingerprint enhancements performs better for good images for an average RMSE equal to 0.1837301 while there is no significant different between the bad and ugly images 0,2228729 and 0.2319876 respectively

Figure 6 presented the RMSE average of good original and enhanced latent fingerprint images.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Homogeneous Subsets for RMSE

Image Category	N	Subset for alpha = 0.05	
		1	2
Good	85	.01837301	
Bad	85		.02228729
Ugly	85		.02319876
Sig.		1.000	.480

Figure 6. Image category mean comparison of RMSE

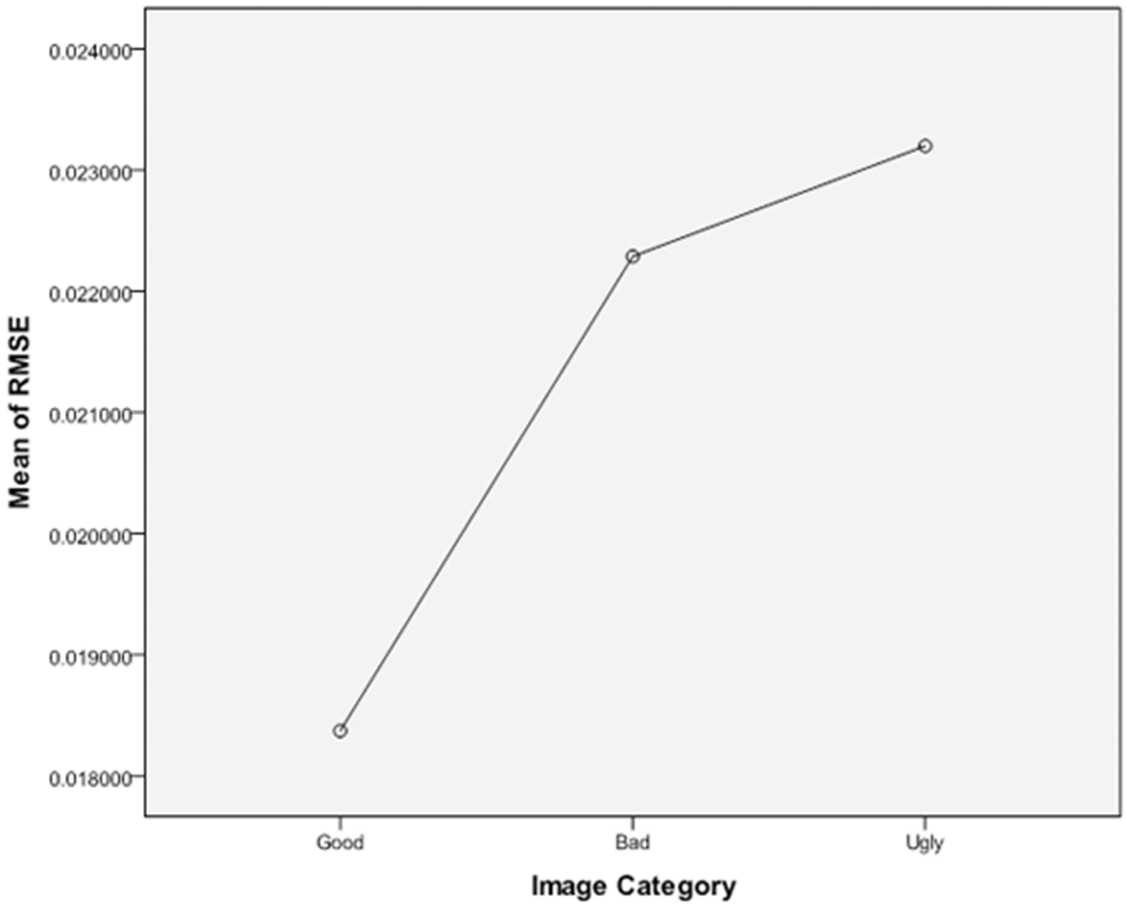


Table 6. RMSE average of Good Images before and after enhancement

Image Category	RMSE Average before enhancement	RMSE average after enhancement
Good	0.048091	0.018373

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6 illustrated RMSE average of good images before and after enhancement, the result shows advantages of proposed model for good images category.

Figure 7. RMSE averages of good images before and after enhancement

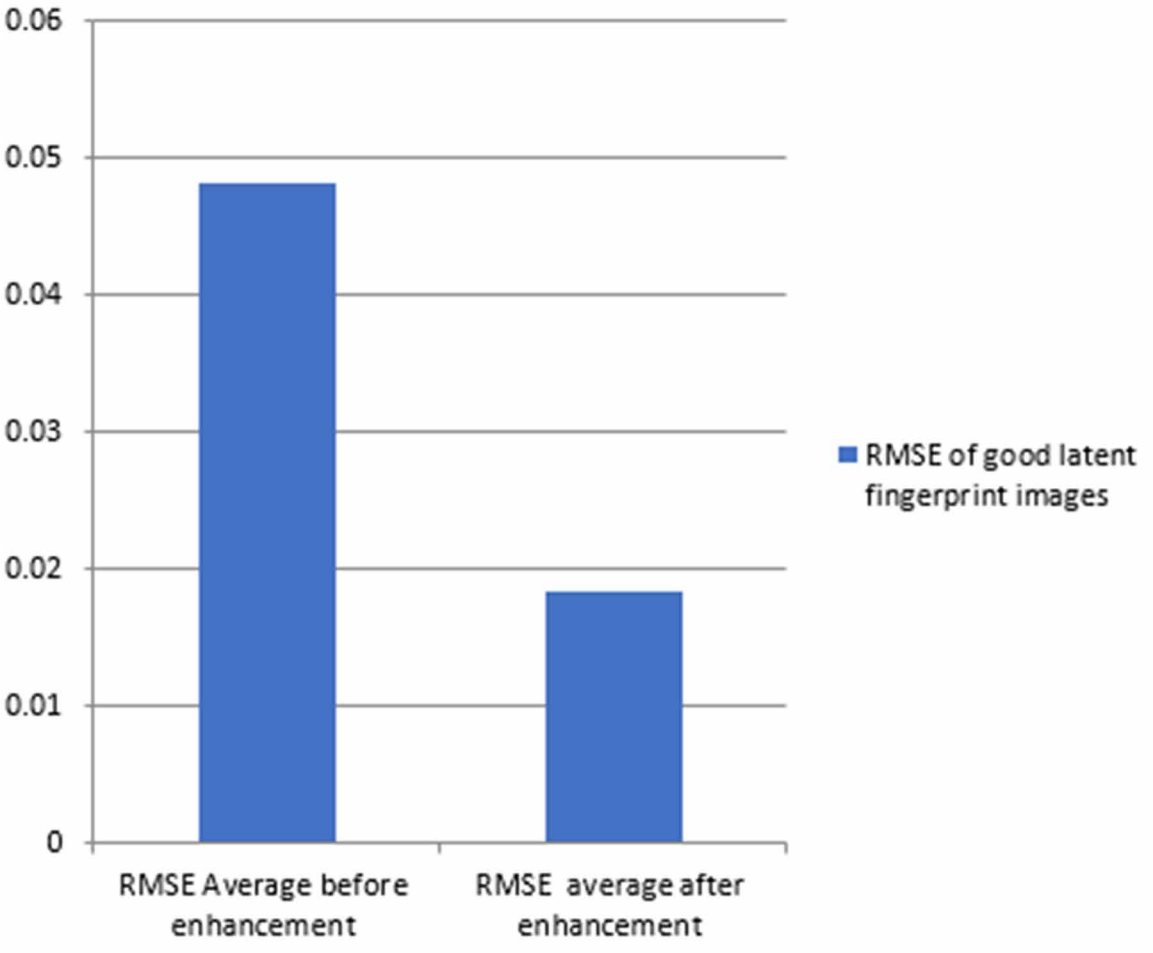
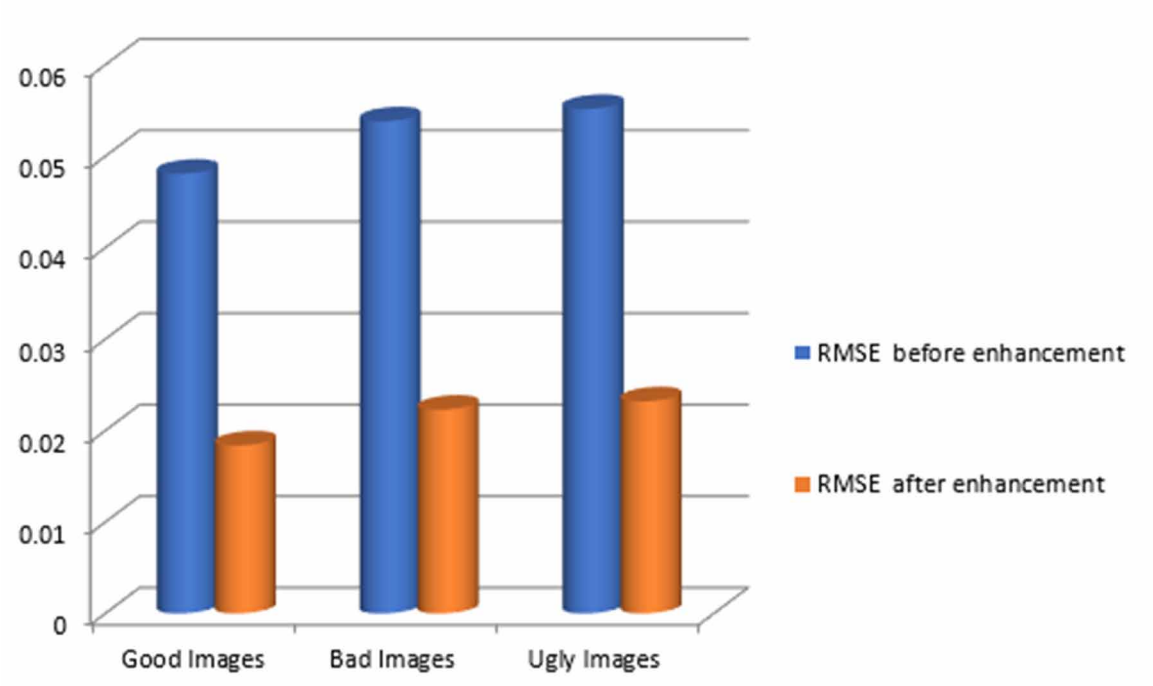


Figure 7 shows the RMSE average of good latent fingerprint after and before enhancement and the comparison for ugly,

To measure the performance ratio of the proposed method, the result after and before enhancement has been compared. Figure 8 presents the RMSE comparison of the different categories before and after enhancement.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. RMSE Image categories comparison



CONCLUSION

Latent fingerprint image quality is an important factor in the performance of minutiae extraction and matching algorithms. Being collected from crime scene, latent fingerprint is often mixed with other components such as structured noise or other fingerprints. The poor image quality of latent fingerprints has a high effect on the accuracy of latent fingerprint Enhancement

Image enhancement is required for latent fingerprint images to improve the accuracy of the proceeding steps. The main goal of a latent fingerprint enhancement is to improve the clarity of the ridge structures while the major challenging problem for latent fingerprint enhancement is to remove various types of image noises while reliably restoring the corrupted regions and enhancing the ridge clarity. Even though considerable progress done in both rolled and plain fingerprint images, latent fingerprint is challenging problem and existing issue in the current research. Rolled and plain fingerprints are obtained in an attended mode so that they are normally of good visual quality while latent fingerprint is collected from crime scenes and usually mixed with other components such as structured noise and other fingerprints. Many of the existing state of the art techniques suffer limitations such as low matching accuracy and high false alarm rate. This is due to the existence of poor-quality images in latent fingerprint with unclear ridge structure and various overlapping patterns together with presence of structured noise.

The chapter has designed latent fingerprint enhancement hybrid technique which is combination of Edge Directional Total Variation Model (EDTV) and quality image enhancement with lost minutia reconstruction techniques based on reliable frequency estimation and orientation field. The objective is to improve the quality of the latent fingerprint image by enhancing the latent fingerprint image.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

To evaluate the result of latent fingerprint enhancement for the proposed technique RMSE, the result after enhancement achieved RMSE average 0.018373, 0.022287, and 0.023199 for the three different image categories available in SD27 data set good, bad and ugly images respectively.

NEW APPROACH FOR LATENT FINGERPRINT ENHANCEMENT

Hybrid technique which is combination of Edge Directional Total Variation Model (EDTV) and quality image enhancement with lost minutia reconstruction techniques based on reliable frequency estimation and orientation field has been proposed to enhance the latent fingerprint images.

FUTURE RESEARCH AND RECOMMENDATION

The effectiveness of the proposed latent fingerprint enhancement presented is highly related to the accuracy of the estimation of frequency and orientation. If the estimation is not reliable it may lead to poor enhancement. Based on the experiment conducted, it's observed that the proposed techniques perform well for the good latent fingerprint images compare to bad and ugly images. Robust techniques are needed for the enhancement of bad and ugly images which may lead to better matching for latent fingerprint. The proposed techniques did not take into account for the overlapped latent fingerprint images. Research on integrating a method that solves the issue of overlapped latent fingerprints would be very beneficial to further improve the proposed solution.

REFERENCES

Cao, K., Liu, E., & Jain, A. K. (2014). Segmentation and Enhancement of Latent Fingerprints: A Coarse to Fine RidgeStructure Dictionary. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(9), 1847–1859. doi:10.1109/TPAMI.2014.2302450 PMID:26352236

Chen, F., Feng, J., Jain, A. K., Zhou, J., & Zhang, J. (2011). Separating overlapped fingerprints. *IEEE Transactions on Information Forensics and Security*, *6*(2), 346–359. doi:10.1109/TIFS.2011.2114345

Feng, J. Z., & Jain, A. K. (2013, April). Orientation field estimation for latent fingerprint enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(4), 925–940. doi:10.1109/TPAMI.2012.155 PMID:22826508

Galbally, J., Haraksim, R., & Beslay, L. (2018). A study of age and ageing in fingerprint biometrics. *IEEE Transactions on Information Forensics and Security*, *14*(5), 1351–1365. doi:10.1109/TIFS.2018.2878160

Gupta, R., Khari, M., Gupta, D., & Crespo, R. G. (2020). Fingerprint image enhancement and reconstruction using the orientation and phase reconstruction. *Information Sciences*, *530*, 201–218. doi:10.1016/j.ins.2020.01.031

Khan, M. A., Khan, T. M., Bailey, D. G., & Kong, Y. (2016). A spatial domain scar removal strategy for fingerprint image enhancement. *Pattern Recognition*, *60*, 258–274. doi:10.1016/j.patcog.2016.05.015

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Krish, R. P., Fierrez, J., Ramos, D., Alonso-Fernandez, F., & Bigun, J. (2019). Improving automated latent fingerprint identification using extended minutia types. *Information Fusion*, *50*, 9–19. doi:10.1016/j.inffus.2018.10.001

Li, J., Feng, J., & Kuo, C. C. J. (2018). Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing Image Communication*, *60*, 52–63. doi:10.1016/j.image.2017.08.010

Liban, A., & Hilles, S. M. (2018, July). Latent Fingerprint Enhancement Based on Directional Total Variation Model with Lost Minutiae Reconstruction. In *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)* (pp. 1-5). IEEE. 10.1109/ICSCEE.2018.8538417

Liu, M., Liu, S., & Yan, W. (2018). Latent fingerprint segmentation based on ridge density and orientation consistency. *Security and Communication Networks*, *2018*, 2018. doi:10.1155/2018/4529652

Liu, S., Liu, M., & Yang, Z. (2017). Sparse coding-based orientation estimation for latent fingerprints. *Pattern Recognition*, *67*, 164–176. doi:10.1016/j.patcog.2017.02.012

Menon, R. R., Sukhadiya, H., & Patel, J. (2015). Standalone USB Flash to USB Flash. *Data Transfer*, ●●●, 6936–6941.

Mukhopadhyay, P., & Chaudhuri, B. B. (2015). A survey of Hough Transform. *Pattern Recognition*, *48*(3), 993–1010. doi:10.1016/j.patcog.2014.08.027

Porwik, P. (2010). *The Modern Techniques of Latent Fingerprint Imaging*. IEEE. doi:10.1109/CISIM.2010.5643695

Qian, P., Li, A., & Liu, M. (2019, June). Latent fingerprint enhancement based on DenseUNet. In *2019 International Conference on Biometrics (ICB)* (pp. 1-6). IEEE.

Revathy, T., Pramila, G., Adhiraja, A., & Askerunisa, A. (2014, April). Automatic Latent Fingerprint Segmentation based on Orientation and Frequency Features. In *2014 International Conference on Communication and Signal Processing* (pp. 1192-1196). IEEE. 10.1109/ICCSP.2014.6950029

Sabhanayagam, T., Venkatesan, V. P., & Senthamaraikannan, K. (2018). A comprehensive survey on various biometric systems. *International Journal of Applied Engineering Research: IJAER*, *13*(5), 2276–2297.

Svoboda, J., Monti, F., & Bronstein, M. M. (2017, October). Generative convolutional networks for latent fingerprint reconstruction. In *2017 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 429-436). IEEE. 10.1109/BTAS.2017.8272727

Yoon, S., Feng, J., & Jain, A. K. (2010). On Latent Fingerprint Enhancement. *Proc. SPIE Biometric Technology for Human Identification VII*.

Zahedi, M., & Ghadi, O. R. (2015). Combining Gabor filter and FFT for fingerprint enhancement based on a regional adaption method and automatic segmentation. *Signal, Image and Video Processing*, *9*(2), 267–275. doi:10.100711760-013-0436-3

Zhang, J. R. L.-C. (2013). Adaptive Directional Total-Variation Model for. IEEE Transactions on Information Forensics And Security, 8.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zhang, J., Lai, R., & Kuo, C. C. J. (2013). Adaptive directional total-variation model for latent fingerprint segmentation. *IEEE Transactions on Information Forensics and Security*, *8*(8), 1261–1273. doi:10.1109/TIFS.2013.2267491

Zhang, R. L., & Kuo, C. J. (2013, August). Adaptive directional total-variation model for latent fingerprint segmentation. *IEEE Transactions on Information Forensics and Security*, *8*(8), 1261–1273. doi:10.1109/TIFS.2013.2267491

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 22
The Business, Societal, and Enterprise Architecture Framework:
An Artificial Intelligence–, Data Sciences–, and Big Data–Based Approach

Antoine Trad
https://orcid.org/0000-0002-4199-6970
IBISTM, France

ABSTRACT

In this chapter, the author presents an artificial intelligence (AI)-based generic concept for decision making using data science. The applied holistic mathematical model for AI (AHMM4AI) focuses on data and access management. A decisive business decision in a business transformation process of a traditional business environment into an automated AI-based business environment is the capacity of the decision-making system and the profile of the business transformation manager (BTM, or simply the manager). The manager and his team are supported by a holistic framework. The role of data science and the needed data modelling techniques are essential for managing data models in a transformation project. For that reason, the development of the big data for AI (BGD4AI) is an essential start.

INTRODUCTION

The success of a *Project* depends on how an Enterprise Architecture for AI (EA4AI), data architecture and modelling activities are synchronized (IMD, 2015).

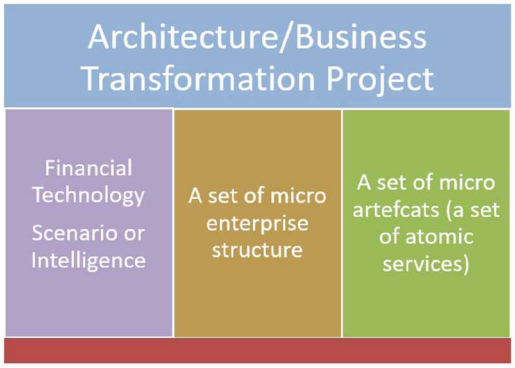
That is why the implementation of such *Project*s requires significant knowledge of data architecture, implementation and modelling techniques. The main GAIP mechanisms are: 1) generic data architecture; 2) implementation interfaces; and 3) data modelling is a part of the Selection management, Architecture-modelling, Control-monitoring, Decision-making, Training management and Project management

DOI: 10.4018/978-1-7998-6985-6.ch022

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Framework's cycles and the data access block



Framework (SmAmCmDmTmPmF, for simplification in further text the term *TRADf* (that stands for the Transformation, Research, Architecture and Development framework) will be used), that supports the *Project*'s activities. As shown in Figure 1, the Data Sciences Integration for Artificial Intelligence (DSI4AI) interacts with all the enterprise's (or simply an Entity) architecture phases, using the data Building Blocks for AI (dBB4AI) or the holistic brick (Trad & Kalpić, 2020a). The chapter is based on complex framework and it is just an extension to it and it includes many subdomains. This is the case of all cross functional and holistic ones. The DSI4AI explains, describes the basics of AI to support the GAIP based transformation project. Such projects are complex undertakings, which are based on the selection and DSI4AI based classification and weightings of the most important critical success factors and areas, which are used as global variables in the author's specialized transformation framework. Where in this chapter the main subject is DSI4AI for GAIP's integration. The GAIP for transformations projects. Such transformation projects can be applied to various Application and Problem Domains (APD), like finance, geopolitics, intelligent cities and other. Such APDs need a specific AI that is based on DSI4AI and AHMM4AI.

BACKGROUND

GAIP and DSI4AI are based on EA4AI, various AI fields, data architecture and modelling technologies. Where the DS4AI should include:

- A data foundational model, or a set of classes/entities, which can be integrated in various architectures; that use calls to various types of algorithms.
- The use of atomic Building Blocks for AI (aBB4AI) concept; which corresponds to an autonomous set of classes.
- The use of a Natural Programming Language for AI (NLP4AI) for development of data interfaces.

The author's long years global research topic's and final Research Question (RQ) (hypothesis #1-1) is: "Which business transformation manager's characteristics and which type of support should be assured for the implementation phase of a business transformation project?" The targeted business domain is any

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

business environment that uses: 1) complex technologies; and 2) frequent transformation iterations. For this phase of research, the sub-question is: "What is the impact of the BGD4AI on EA4AI and *Projects*?"

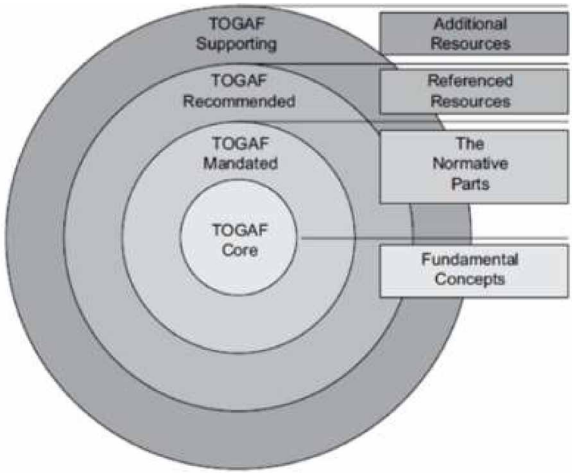
MAIN FOCUS OF THE CHAPTER

In this research phase the author's target is BGI4AI's that is a part of the Architecture module (Am), and he tries to prove that such a concept can be built on a loosely coupled architecture. The BGI4AI uses the Data Management Concepts for Artificial Intelligence (DMC4AI) to interface data sources. The DS4AI uses Mathematics for Data Science or the already presented AHMM4AI, which deals with mathematical models and algorithms that are used to analyse data, offers conclusions and it supports the DS4AI. *Projects* are increasingly digital and data is global; these huge amounts of data are full of valuable operational information. The BGI4AI supports data extraction in a way that to be used by AHMM4AI. In this chapter the AHMM4AI uses algorithms which are essential for analysing data and thus providing a basis for evidence-based decisions in various domains and in complex situations.

Data Management and Complexity

Chaotic and complex DSI4AIs can cause failures, where business environments have complex data architectures, as shown in Figure 2 (IBM, 2015a). It is useful to interface exiting EA4AI structures and resources, like the use of documents, as shown in Figure 2.

Figure 2. Structure of an architecture document (The Open Group, 2011)



DS4AI Basics

The main DS4AI basics are (Guru99, 2021):

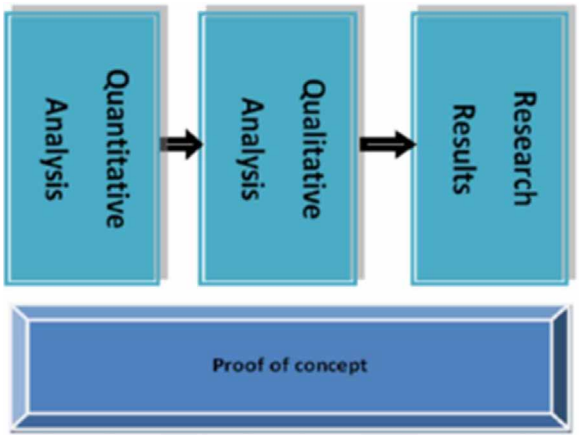
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- It involves extracting insights in the context of vast amounts of data, by using scientific methods, algorithms and processes. It helps in finding hidden patterns from raw data. DS4AI is the result of evolution of statistics, data analysis, and the use of BGD4AI.
- It is cross functional and extracts knowledge from structured or unstructured databases. DS4AI translates business problems into a research project and then translates it into solutions.
- It involves extracting insights from vast amounts of data by the use of various scientific methods, algorithms and processes.
- Includes, statistics, visualization, Deep Learning Integration for AI (DLI4AI), Machine Learning Integration for AI (MLI4AI) concepts.
- Its Process goes through discovery, data preparation, data models' planning, model building, operation management and the communication of results.
- The actual possible roles are: 1) Data scientist; 2) Data engineer; 3) Data analyst; 4) Statistician; 5) Data architect; 6) Data admin; 7) Business analyst; and 8) Data/analytics manager.
- Tools that can be used, are: R, SQL, Python, SaS…
- Predictions of Business Intelligence, is looking backward (history), whereas DS4AI looks forward (actual and future).
- Possible applications are 1) Internet search; 2) Recommendation systems; 3) Image and speech recognition; 4) Gaming world, like eSports; 5) Online price comparison; and 6) many other…
- Various information, data formats and sources are the biggest challenges for DSI4AI and BGI4AI.

THE RESEARCH DEVELOPMENT PROCESS

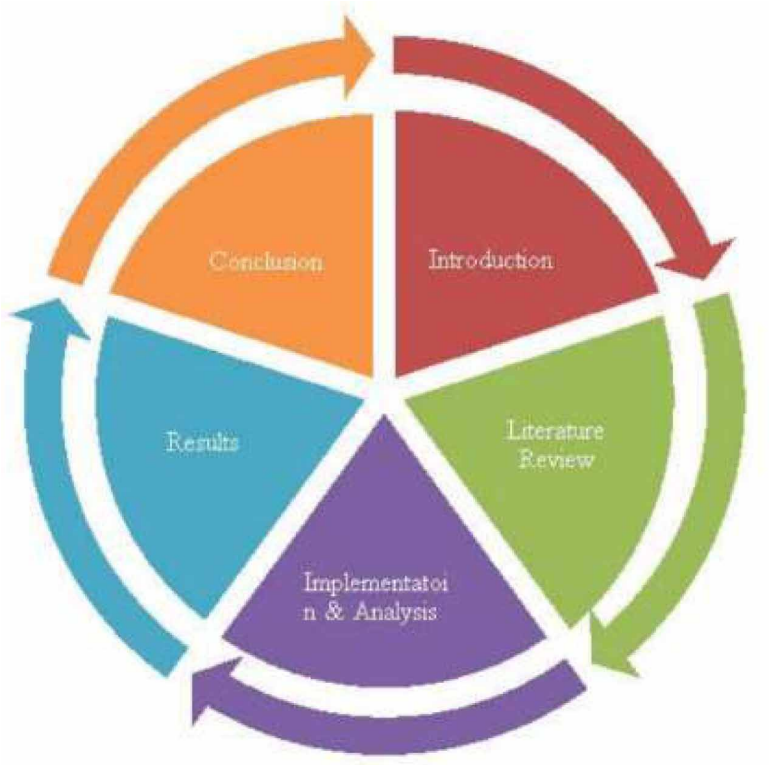
As shown in Figure 3, the Research and Development Process for AI (RDP4AI) focuses on the impacts of the mechanistic BGI4AI's integration and uses a mixed hyper-heuristics based methodology (Vella, Corne, & Murphy, 2009). The RDP4AI is based on an extensive cross-functional Literature Research Process for AI (LRP4AI), a Qualitative Analysis for Artificial Analysis (QLA4AI) methodology and on a Proof of Concept (PoC)for the proposed hypotheses.

Figure 3. The mixed method flow diagram (Trad & Kalpić, 2020a)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

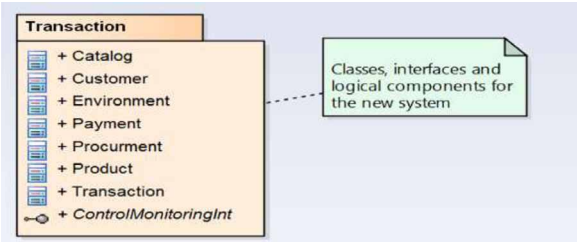
Figure 4. The experiment's overall diagram of components



Projects use a PoC or prototyping to establish BGI4AI's: 1) feasibility; 2) viability; 2) major technical issues; and 3) offer recommendations (Camarinha-Matos, 2012).

The PoC is used to prove the feasibility of the BGI4AI, as shown in Figure 4. The PoC is based on a transaction management case that uses data class diagram for a data transaction, as shown in Figure 5; where an DS4AI module is called.

Figure 5. The experiment's class diagram package



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Critical Success Areas, Factors and DSI4AI

Critical Success Area (CSA) is a selected set of Critical Success Factors (CSF), where the CSF is a set of Key Performance Indicators (KPI), where each KPI corresponds/maps to a single *Project* requirement and/or problem type. For a given requirement (or problem), an *Entity* architect can identify the initial set of CSAs and their CSFs to be managed by the Decision-Making System for AI (DMS4AI). Hence the CSFs are important for the mapping between the problem types, knowledge constructs, organisational items. Therefore, CSFs reflect possible problem types that must meet strategic *Project* goals and predefined constraints. Measurements are used to evaluate performance in each of the CSA sets, where CSFs (and KPIs) can be internal or external to the environment; like: 1) problem type or gap analysis is an internal CSF; and 2) DMS4AI in real time and in minimal time is also an internal one. Once the initial set of CSFs has been identified, then the *Project* can use the DMS4AI to propose a set of solution types. The proposed BGI4AI based DSI4AI delivers a set of solution instance and recommendations (Trad & Kalpić, 2020a).

THE TARGETED DOMAIN

Main Stages

This chapter's targeted domain characteristics are (OMNI-SCI, 2021):

- Definition: DS4AI includes the mining of large data sets of raw data, both structured and unstructured, and is used to identify patterns and extract actionable insight from them. It includes, statistics, inference, computer science, predictive analytics, MLI4AI and BGD4AI.
- Lifecycle: The 1st stage in DS4AI's pipeline workflow involves, acquiring data, extracting data, and entering data into the system. The 2nd stage is the maintenance stage, which includes Data Warehouses Integration for AI (DWI4AI), data cleansing, data processing, data staging and data architecture activities.
- Data processing follows is a fundamental activity that includes, data exploration and analytics that DS4AI specialists manage. It involves also: data mining, data classification and clustering, data modelling and summarizing insights from the data.
- The data analysis stage includes exploratory and confirmatory tasks, regression, predictive analysis, QLA4AI and text mining.
- During the final stage, the data scientist communicates the found insights; and this stage involves data visualization, data reporting, the use of various Business Intelligence Integration for AI (BII4AI) tools, and assisting businesses, policymakers and others activities.

DSI4AI Preparation and Exploration

DSI4AI's preparation and exploration main characteristics are (OMNI-SCI, 2021):

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Data preparation and analysis are important DS4AI activities, but data preparation consumes 60 to 70 percent of the DS4AI activities; rarely generated data is in a correct, structured and noiseless form. In this phase, the data is transformed for further use.
- This process involves: 1) The transformation and sampling of data; 2) Checking both the features and observations; and 3) Using statistical techniques to remove noise. This step also illuminates whether various features in data sets are independent and whether there are incorrect values.
- This exploration phase is the principal difference between DS4AI and Data Analytics for AI (DA4AI), where the DS4AI has a macro view, aiming to propose precise questions on data to extract and to deliver insights and knowledge from this data. DA4AI contains the question(s) and the results.

DSI4AI Modelling

DSI4AI modelling characteristics are (OMNI-SCI, 2021):

- It fits data into the model using MLI4AI algorithms, where the model's selection depends on the type of data and the business requirement.
- Next, the model is tested for accuracy and other constraints that enables the adjustment of the model and to achieve acceptable results. If the model does not satisfy the requirements, another DS4AI model can be used.
- When testing succeeds with good data, it produces the desired results for the BI requirement, the DS4AI model can be finalized and deployed.

The Importance of DS4AI

Actually, there are about 40 zettabytes of used data and it is growing exponentially. Internet users generate about 2.5 quintillion bytes of data every day. Every person on Earth, is generating about 146,880 GB of data every day and by the year 2025, 165 zettabytes will be used every year. DS4AI tools are critical to understand BGD4AI and the used data from multiple sources (OMNI-SCI, 2021).

DS4AI and DA4AI

DS4AI and DA4AI are different; where the DS4AI starts earlier, exploring massive data sets, investigating their potential, identifying trends and insights and visualizing the results; and DA4AI comes in at a later stage. These processes report the results, make prescriptions for improving performances based on the analysis processes and optimize the use of data tools. DA4AI analyses a specific dataset of structured or numerical data using a given research question. BGI4AI manages larger masses of both structured and unstructured data. DA4AI uses historical data in a context and has less to do with predictive modelling. DA4AI needs precise questions, where the DS4AI does not create statistical models or use MLI4AI tools. DA4AI focuses on evaluating strategies for businesses, by comparing data assets to various organizational hypotheses. DA4AI works with localized data that has already been processed. DA4AI and DS4AI need mathematical, analytical and statistical skills, where the DS4AI is more predictive and the relationship between the DA4AI and data is retrospective (OMNI-SCI, 2021).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

RDP4AI's CSF's

The RDP4AI's results are shown in Table 1.

Table 1. The factors have an average of 9.25

Critical Success Factors	KPIs	Weightings
CSF_DS4AI_LRP4AI	Proven	From 1 to 10. **10 Selected**
CSF_DS4AI_PoC	Feasible	From 1 to 10. **09 Selected**
CSF_DS4AI_References	Feasible	From 1 to 10. **09 Selected**
CSF_DS4AI_EA4AI	Proven	From 1 to 10. **10 Selected**
CSF_DS4AI_CSF_CSA	Feasible	From 1 to 10. **09 Selected**
CSF_DS4AI_Mixed_Methods	Feasible	From 1 to 10. **09 Selected**
CSF_DS4AI_Transformation_Setup	Feasible	From 1 to 10. **09 Selected**
CSF_DS4AI_DMS4AI_Interfacing	Feasible	From 1 to 10. **09 Selected**

valuation

As shown in Table 1, the result's aim is to prove or justify the RDP4AI's feasibility; and the result permits to move to the next CSA that is the AHMM4AI.

AHMM4AI'S SUPPORT FOR DS4AI

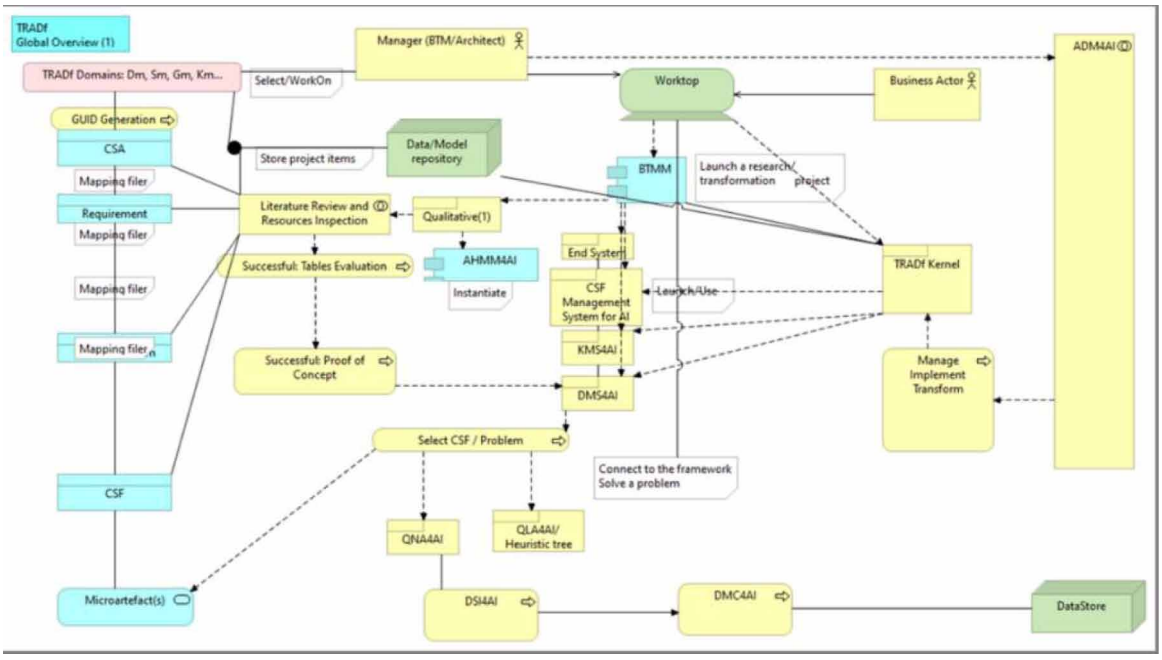
TRADf offers an NLP4AI that can be used to describe and implement the GAIP, BGI4AI, DMS4AI, DSI4AI, QNA4AI and QLA4AI based scenario(s), as shown in Figure 6, which can be used to solve *Project* requests. It is a complex domain and it requires the knowledge and experiences in practically all AI and ICS fields.

A Generic Holistic Approach

The *TRADf* proposes a holistic approach to analyse events and eventually manage possible risks to help Application and Problem Domains (APD) Managers to avoid major pitfalls. Traditionally, complex risk concepts, were associated with a single origin or CSF; mainly personified to concretise a complex situation. Pitfalls may be defined as a violation of an internal risk's related CSF that can be due to various types of problems or constraints; where in DSI4AI can be applied for example, for: 1) estimating

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The overview



Project risks; 2) APD risk initiatives; 3) special domains, like geopolitical events or telecommunication implementations; 4) tracking financial crimes; and 3) other CSFs...

The Microartefacts' Distributed Architecture Model for the GAIP

The AHMM4AI has a dynamic defined nomenclature to facilitate its integration in any *Entity* EA4AI model, like the Architecture Development Method for AI (ADM4AI). The AHMM4AI is the *Entity's* holistic structural model and is a set of multiple coordinated DSI4AI processing to deliver solutions

Figure 7. The applied model's basics nomenclature (Trad & Kalpić, 2020a)

Basic MM's Nomenclature

Iteration	= An integer variable that denotes a *Project/ADM iteration*	
microRequirement	= KPI	(1)
CSF	= Σ KPI	(2)
CSA	= Σ CSF	(3)
Requirement	= ⋃ microRequirement	(4)
microKnowledgeArtefact	= ⋃ knowledgeItem(s)	(4)
neuron	= action ° data + microKnowledgeArtefact	(5)
microArtefact	= ⋃ (e)neurons	(6)
microEntity or Enterprise	= ⋃ microArtefact	(7)
Entity or Enterprise	= ⋃ microEntity	(8)
microArtefactScenario	= ⋃ microArtefactDecisionMaking	(9)
Decision Making/Intelligence	= ⋃ microArtefactScenario	(10)
EnityIntelligence	= ⋃ Decision Making/InteligenceComponent	(11)
MM(*Iteration*) as an instance	= EnityIntelligence(*Iteration*)	(12)

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

that correspond to various just in time processing schemes which use the same *Project*'s central pool of CSAs and CSFs. The basic AHMM4AI nomenclature that the base for DSI4AI, is presented in Figure 7, to the reader in a simplified form, to be easily understood on the cost of a holistic formulation of the AHMM4AI's basics for QNA4AI and QLA4AI. The DMS4AI uses an AHMM4AI's instance to solve a *Project* problem.

The proposed *Entity's* AHMM4AI instances enables the possibility to support the BGI4AI based DSI4AI; using CSFs weightings and ratings (in phase 1), and based on multicriteria evaluation (selected and defined constraints). The symbol å indicates summation of all the relevant named set members, while the indices and the set cardinality have been omitted. The proposed AHMM4AI should be understood in a broader sense, more like set unions. As shown in Figure 7:

- The abbreviation "mc" can be used, and stands for micro, which depends on the granularity.
- The symbol å indicates summation of weightings/ratings, denoting the relative importance of the set members selected as relevant. Weightings as integers range in ascending importance from 1 to 10 (or another range defined by GAIP based analysists).
- The symbol $\underline{\cup}$ indicates sets union.
- The proposed AHMM4AI enables the possibility to define *Project* DSI4AI as a model; using CSFs weightings and ratings evaluation.
- The selected corresponding weightings to: CSF ϵ { 1 … 10 }; are integer values, that are presented in tables. The rules were presented in the RDP4AI section.
- The selected corresponding ratings to: CSF ϵ { 0.00% … 100.00% } are floating point percentage values.

The AHMM4AI's Structure for DSI4AI Based Solutions

The AHMM4AI's has a composite structure that can be viewed as follows:

- The static view has a similar static structure like the relational model's structure that includes sets of CSAs/CSFs that map to tables and the ability to create them and apply actions on these tables (Lockwood, 1999).
- In the behavioural view, these actions are designed using a set of AHMM4AI nomenclature, the implementation of the AHMM4AI is in the internal scripting language, used also to tune the CSFs (Lazar, Motogna, & Parv, 2010).
- The skeleton of the *TRADf* uses microartefacts' scenarios to support just-in-time DSI4AI requests.

Entity/Enterprise Architect as an Applied Mathematical Model

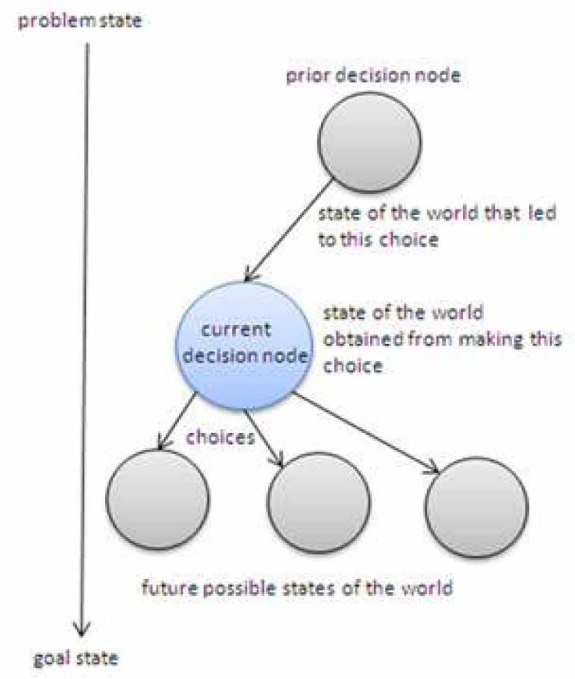
A generic *Entity* architecture model and its ADM4AI are the kernel of this RDP4AI and they are the basics of its *TRADf*. The author wants to propose the AHMM4AI to represent the GAIP. The LRP4AI has shown that existing LRP4AI's resources on DSI4AI, are practically inexistent. This pioneering research work is cross-functional and links all the DSI4AI or QLA4I/QNA4AI based microartefacts to an *Entity*; where the main reasoning component is a DSI4AI engine that is based on heuristics.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Heuristics, Empirics and Action Research

The DSI4AI is based on a set of synchronized AHMM4AI instances, where each AHMM4AI can launch a QLA4AI beam-search based heuristic processing (Kim, & Kim, 1999; Della Croce & T'kindt, 2002). Weightings and ratings concept support the AHMM4AI to process a GAIP request for an optimal analysis or solution for a given *Project* problem. Actions Research for AI (AR4AI) (Berger, & Rose, 2015) can be considered as a set of continuous beam-search heuristics processing phases and is similar to design, analysis and architecture processes, like the ADM4AI (Järvinen, 2007). Fast changing *Entity's* change requests may provoke an important set of events and problems that can be hard to predict and solve; that makes the GAIP various types of actions useless and complex to implement. The AHMM4AI is responsible for the DSI4AI that uses QLA4AI heuristic process for *Entity*'s problem solving and synchronizes a set of AHMM4AI instances which have also separate heuristics processes and are supported by a dynamic tree algorithm, as shown in Figure 8 (Nijboer, Morin, Carmien, Koene, Leon, & Hoffman, 2009) that manages tree nodes and their correlation with memorized patterns that are combinations of data states and heuristic goal functions. The AHMM4AI capacities are measured by analysing the *TRADf's* AHMM4AI tree.

Figure 8. The applied heuristics tree algorithm (Nijboer, Morin, Carmien, Koene, Leon & Hoffman, 2009)



Holistic AR4AI based heuristics enables reflective practice that is the basis of a holistic approach to develop EA4AI based GAIP solutions, where its kernel and skeleton are a dynamic DMS4AI (Leitch, & Day, 2006). GAIP is based on DSI4AI, QLA4AI and QNA4AI methods (Loginovskiy, Dranko, & Hollay, 2018).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Qualitative, Quantitative (or Mixed) Related to the Notion of Time, Space and Scope

As already mentioned, the AHMM4AI and its underlining set of created instances is mainly a QLA4AI beam-search heuristic tree (Della Croce, & T'kindt, 2002). In each of the tree's node a precise call to DSI4AI functions (or other) can be executed, by precision or objectivity the author refers to inputted data, constraint (and/or rules) and above all: 1) Time related to timestamp tracing system; 2) Space, related to the analysed *Entity's* space; and 3) Scope of GAIP's process. These facts, form the basis of an applicable GAIP and DSI4AI are based on AHMM4AI instances.

The Applied GAIP Transformation Mathematical Model

GAIP is a part of the *TRADf* that uses microartefacts to support just-in-time DMS4AI actions. The GAIP based component and interface, are based on a light version of the ADM4AI, having a systemic approach. A *Project* using GAIP is the combination of GAIP based, EA4AI methodology (like the TOGAF's ADM4AI) and the proposed AHMM4AI, that is presented in Figures 9 and 10.

Figure 9. The model's generic structure

The Generic AHMM's Formulation

$$AHMM = \bigcup ADMs + MMs \quad (13)$$

The generic AHMM can applied to any specific domain; in this chapter and RDP4AI's phase, the *Domain,* is GAIP based and the AHMM4AI = AHMM(APD), as shown in Figure 10; where AHMM can applied to any domain and any concept.

Figure 10. The AHMM4AI structure

AHMM's Application and Instantiation for a Specific Domain

$$Domain = \text{Geopolitical Analysis (GA)} \quad (14)$$

$$AHMM(Domain) = \bigcup ADMs + MMs(Domain) \quad (15)$$

The proposed combination can be modelled after the following formula for the GAIP Transformation Mathematical Model (GAIPTMM) that abstracts the *Project* for a given *Entity*:

(AHMM4AI for an *Iteration*):

iAHMM4AI = AHMM4AI(*Iteration*);

$iAHMM4AI = Weigthing_1 * iAHMM4AI_Qualitative + Weigthing_2 * iAHMM4AI_Quantitative$ (16).

The *Project's* AHMM4AI (PAHMM4AI) = å iAHMM4AI for an ADM4AI's instance (17).

(GAIPTMM):

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

GAIPTMM = å PAHMM4AI instances (18).

The Main Objective Function for GAIP based (MOFGAIP) of the GAIPTMM's formula can be optimized by using constraints and with extra variables that need to be tuned using the AHMM4AI. The variable for maximization or minimization can be, for example, the *Project* success, costs or other (Dantzig, 1949). For this chapter's PoC the success will be the main and only constraint and success is quantified as a binary 0 or 1. Where the MOF4AI definition will be:

Minimize risk GAIPTMM Function (GAIPTMMf) (19).

The AHMM4AI is based on a concurrent and synchronized *TRADf*, which uses concurrent threads that can make various AHMM4AI instances run in parallel and manage information through the use of the AHMM4AI's NLP4AI. The GAIPTMM is the combination of the GAIP, *Project* and EA4AI methodologies and a holistic AHMM4AI that integrates the *Entity* or organisational concept, Information and Communication Systems (ICS) that have to be formalized using a functional development environment like *TRADf's* NLP4AI.

The AHMM4AI's CSFs

Based on the LRP4AI, the most important AHMM4AI's CSFs that are used are evaluated to the following:

Table 2. The factors have an average of 10.0

Critical Success Factors	KPIs	Weightings
CSF_DSI4AI_AHMMAI_TRADf_Integration	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_InitialPhase	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_PoC	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_Qualitative&Quantitative	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_Final_Instance	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_ADM4AI_Integration	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_AHMMAI_APD_Interfacing	Complex	From 1 to 10. **08 Selected**

valuation

Table is wider than 6.25 inches

As shown in Table 2, the result's aim is to prove or justify that it is complex but possible to implement atheAHMM4AI in the *Entity's* ICS. The next CSA to be analysed is DMC4AI as an interface.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

DMC4AI AS AN INTERFACE

DMC4AI Based DSI4AI Integration Strategy

The DSI4AI strategy should be based on well-defined principles, like (IBM, 2015a):

- Data have to be easy to enter and must be accurate.
- Data collections are based on data modelling concepts.
- Data must be available to business personnel.
- Business processes for data management have to be automated.
- Data should not be redundant.
- Data have to be recorded accurately and must be auditable.
- The cost of data collection procedures have to be optimized.
- DSI4AI must adopt standards for common data models.

Data Transformation Models

In DSI4AI, Data transformation is the process of transforming data from one format or class into another format or class. It is an essential process for data integration and data management tasks, like data wrangling, data warehousing, data integration and EA4AI.

The Meta-Model

This PoC uses a meta-model to show the relationships between the *Project*'s different types of data models and data sources. This *Project*'s meta-model contains descriptions on how data sources are handled in the newly transformed business system's and of the types of used data structures

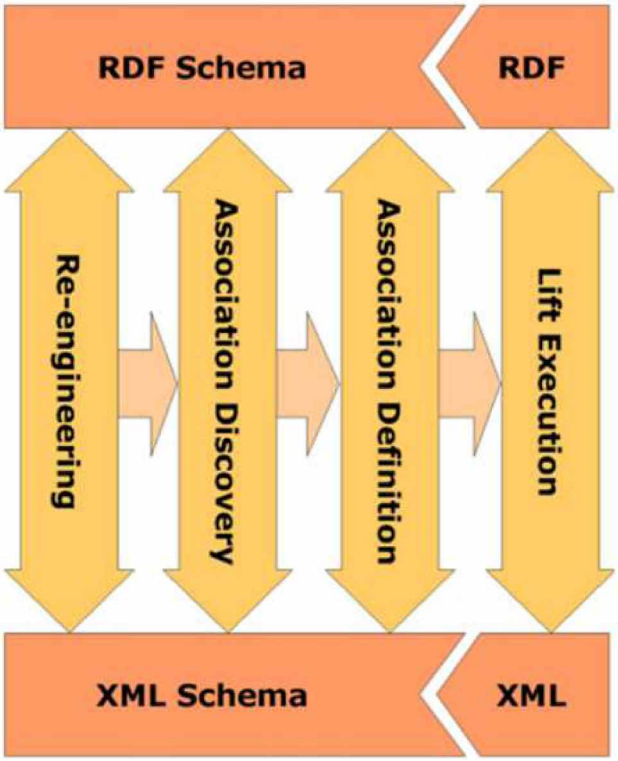
NORMALIZATION AND INTER-OPERABILITY

eXtensible Markup Language (XML)

DSI4AI must be based on various standards that are all in turn based on XML technologies. XML enables ontology-mediated interoperability that is based on data schemas. Schemas enable the alignment by defining common data structures. Data architecture and modelling conceptual schemas are used in the mediation module that describes the semantics of data sources and destinations. Data architecture and modelling conceptual normalisation is a three-layered approach for information representation, as shown in Figure 11, the three layers for data integration include: 1) syntax checking; 2) data modelling; and 3) semantics control. The PoC uses the normalisation conversion modules that include a set of tables to automate the data translation process (Fodor, Dell'Erba, Ricci, Spada, & Werthner, 2002).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Stages of conceptual normalisation process (Fodor, Dell'Erba, Ricci, Spada, & Werthner, 2002)



JavaScript Object Notation

JavaScript Object Notation (JSON) is an open standard file format used as a data interchange format, that has a human-readable text to store and manage data objects. These objects have attributes in the form of value pairs and array data types. It is an easy and common data format, with a diverse range of applications. JSON is replacing XML because it is a language-independent data format and is derived from JavaScript.

The Model's Unit of Work

A holistic alignment, identification and classification of all the RDP4AI's resources must be done, so that the research process can start. A holistic alignment needs also, to define the Unit of Work (UoW) or the basic microartefact. Using the "1:1" mapping concept, the microartefact is represented with a class diagram and can be represented also by an XML model. Such a mapping concept, is based on an automated naming convention that can link all the RDP4AI's resources. The mapping concept supports the interoperability between all the RDP4AI's modules and enables the use of NLP4AI microartefacts that uses ADM4AI and DSI4AI (Mehra, Grundy, & Hosking, 2005).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

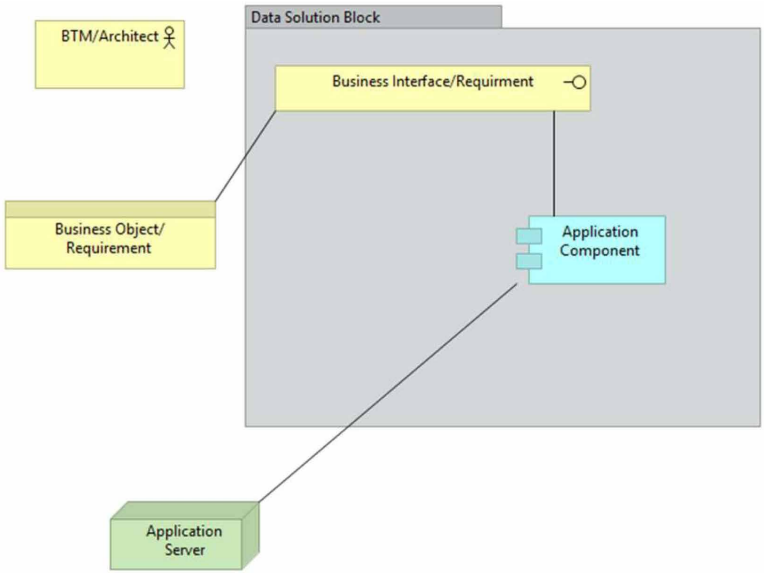
The Object Relational Mapping

Object Relational Mapping (ORM) in ICS is a programming technique for transforming data between various types of systems. ORM uses OO programming languages like, C++, Java, C# and many other. ORM creates a *virtual object database* that is used by programming language.

Data Solution Blocks

Upon a concrete *Project* requirement, the *Manager* issues a contract to resolve this requirement by using EA4AI. EA4AI ensures that new requirements are managed accordingly to the *Project's* records and objectives. The requirement is linked to an instance of a newly created data Building Block for AI (dBB4AI) and its instance (a data Solution Block for AI, dSB4AI). The dBB4AI is a part of GAIP; and as shown in Figure 12, the *TRADf* uses the GAIP and DSI4AI that include the patterns needed to integrate dSB4AIs.

Figure 12. The experiment's data block diagram



Transactions Data Modelling

The *Project* must put important considerations on technical assertion, governing, control, access management and monitoring of business transactions that are based on the Model View Controller for AI (MVC4AI) pattern. Business transactions use the following standards: 1) the policy framework; and 2) business messaging that is needed for domain-specific integration (OASIS, 2009).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Data Lifecycle Diagram

The Data Lifecycle Diagram (DLD), manages DSI4AI's data throughout their lifecycle, from EA4AI, modelling, processing to removal, within the defined rules and constraints. The data sets are loosely coupled with DSI4AI's processes and that allows the definition of common data models enabling an effective data and other resources' sharing. Each data model must have its status that contains the transitions and the gaps between phases. A phase's state must contain a stable data model, where the data EA4AI and modelling lifecycle enables a qualitative design of the related business entities. The resultant data models must be atomic or simple in order to (re)engineer the *Project*'s business processes (TOGAF Modelling, 2015c).

DMC4AI's CSF's

Based on the business case's (and its CSA) LRP4AI process managed and weighted the most important CSFs that were used.

Table 3. The factors have an average of 9.40

Critical Success Factors	KPIs	Weightings
CSF_DMC4AI_Interface_Strategy	Feasible	From 1 to 10. **09 Selected**
CSF_DMC4AI_Interface_Transformation	Proven	From 1 to 10. **10 Selected**
CSF_DMC4AI_Interface_Mapping	Proven	From 1 to 10. **10 Selected**
CSF_DMC4AI_Interface_EA4AI	Proven	From 1 to 10. **10 Selected**
CSF_DMC4AI_Interface_CSF_CSA	Feasible	From 1 to 10. **09 Selected**
CSF_DMC4AI_Interface_Modelling	Complex	From 1 to 10. **08 Selected**
CSF_DMC4AI_Interface_DLD	Proven	From 1 to 10. **10 Selected**
CSF_DMC4AI_Interface_Plaform	Feasible	From 1 to 10. **09 Selected**

valuation

As shown in Table 3, the result's aim is to prove or justify the RDP4AI's feasibility; and the result permits to move to the next CSA that is the data access, architecture and agility.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

DS4AI PLATFORM

Basic Roadmap

The major notions to implement a DSI4AI oriented platform are (Veneberg, 2014);

- EA4AI contains a traceability and monitoring subsystem. Combining DMS4AI and DSI4AI consolidates the *Entity's* support for business specialists, without the need to trace the used data. The EA4AI provides a meta-data model for operational and historic data that are needed for DMS4AI routine operations.
- For BGI4AI, Data Base Management System (DBMS) is implemented to access, read and write data sets.
- There are three types of DBMS: 1) navigational DBMS; 2) Relational DBMS (RDBMS); and 3) the Post-RDBMS (PRDBMS). RDBMS and PRDBMS are the most popular ones and the Simple Query Language (SQL) is used with RDBMS, that manipulates tables.
- Data is stored in two-dimensional racks and it can be accessed, by read and write transactions using a RDBMS.
- There is also non-relational database, labelled No Simple Query Language (NoSQL) Databases (NoSQLDB), that are the successors of RDBMS. RDBMS focuses on its relations between tables, where the NoSQLDB focuses on its objects' models.
- NoSQLDB using key-value stores, are optimal for DSI4AI which processes large volumes of data; and NoSQLDB benefits are latency and throughput. This is the main fact why many AI applications like BGD4AI and performant real-time web applications NoSQLDBs.
- NoSQLDB and RDBMS are two-dimensional data structure or a multidimensional data structure; where a one-dimensional data structure is an array of items. Data items are stored in central memory as a sequence of items, starting at row 0 to row n.
- Data structures that are composed of more than one dimension are called multi-dimensional data structures (or cubes).
- DWI4AI is used to store operational and historical data that can be used as a reference source for DSI4AI in support for the DMS4AI.
- DWI4AIs lacks data traceability and monitoring, where a *Project* needs instructions on how to trace data sets and events.
- DWI4AI is a decision support module for executives, analysts and architects.
- Extract Transform and Load (ETL) has the following characteristics:
    - Supports the extraction of data from legacy systems and transformed into *Entity* data.
    - Has many functions, that includes logical conversion of data, domain verification, conversion from one RDBMS to another.
    - It supports the: 1) creation of default values; 2) summarization of data; 3) addition of time values to the data key; 4) restructuring of the data key; 5) merging of records and deletion of extraneous or redundant data.
    - It helps in transforming data coming from operational source systems into *Entity* data to be accessed by DWI4AI. The transformation of data is usually processed in an Operational Data Store (ODS).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- The ODS the data updated from an OnLine Transaction Processing (OLTP) response time; where the ODS is a hybrid environment in which *Project* data is transformed by an ETL into a defined format. The ODS abstracts raw data DWI4AI from *Entity*.
- Integration processes support the DSI4AI to become a consistent source package.

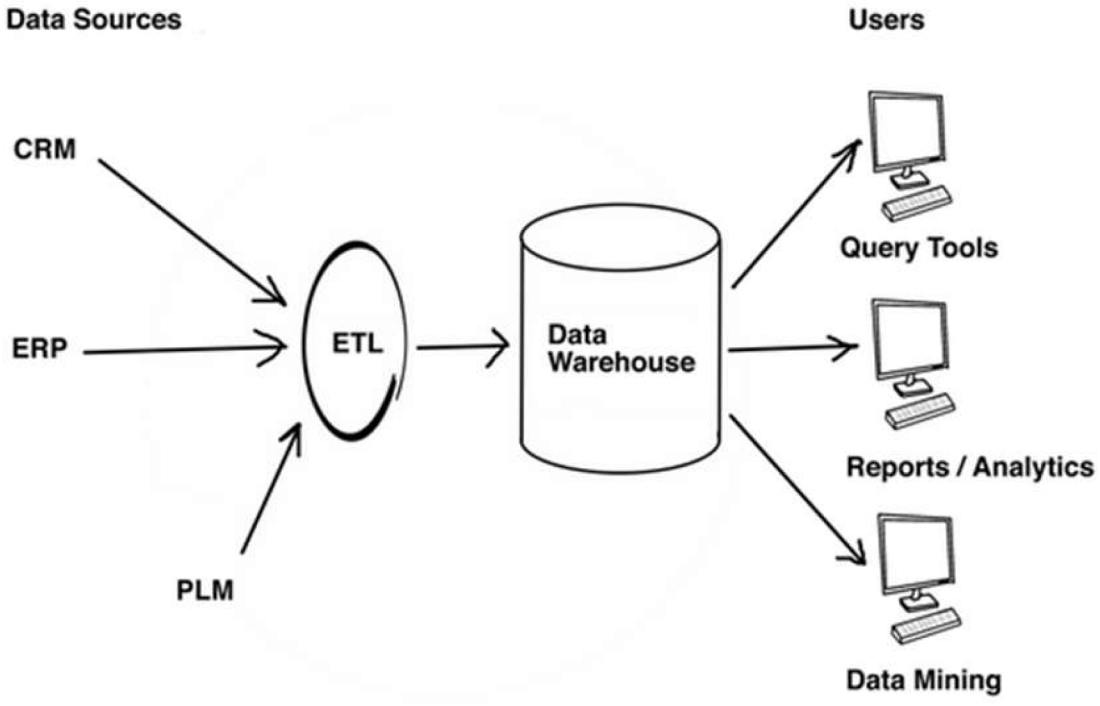
Enterprise Service Bus and Enterprise Application Integration

An Enterprise Service Bus (ESB) implements an enterprise-wide communication subsystem between mutually interacting applications using the Service Oriented Architecture (SOA) paradigm. It represents a communication architecture for global distributed computing, and is a special variant of the more general client server concept, where *Entity* applications and modules may be a server or client. An ESB promotes modularity, agility and flexibility in relation to the ICS and *Project*. It is used in Enterprise Application Integration (EAI) approach for heterogeneous and complex services-based *Projects*. *Project*s must use ESB to glue the various data sources of the business environment, through the use of the technology stack and data connectors, which permit a holistic data services' management.

Extraction, Transform and Load

ETL processes are defined as accessing data stored in various locations and transforming them in order to enable their unification, quality or normalization. The DSI4AI, proposes the separation of data processing activities; that enables data services to access data without bothering about various data sources'

Figure 13. Various data sources (Tamr, 2014)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

complexities. The DSI4AI insures: 1) intra (or extra) data transparency; 2) managing accessibility; and 3) data quality control (Tamr, 2014). As shown in Figure 13, the ETL processes are responsible for the access of data from heterogeneous data sources. The challenge lays in the real-time data transformation and normalization processes (Trujillo & Luján-Mora, 2003).

BGD4AI for DSI4AI

BGD4AI has the following characteristics (SAS, 2021):

- BGD4AI refers to data that is massive, fast and is very complex to process using existing ICS' resources.
- Data for DSI4AI can managed by the BGD4AI that is based on the three V's: 1) Volume; 2) Velocity; and 3) Variety.
- Volume: *Entities* collect data from a variety of sources, including business transactions, smart devices, industrial equipment, videos, social media and more.
- Velocity: The evolution of ICS' and data streams with extreme speed, requests just-in-time processing. This needs performance and scalability.
- Variety: *Entity's* data have various types of formats, from structured, numeric data in traditional databases to unstructured text documents, emails, videos... That needs a unique transformation platform.

Table 4. The factors have an average of 9.0

Critical Success Factors	KPIs	Weightings
CSF_DSI4AI_Platform_Roadmap	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_Platform_Interface	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_Platform_Legacy_Mainframes	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_Platform_EA4AI_ADM4AI	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_Platform_CSF_CSA	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_Platform_Scalability	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_Platform_ORM	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_Platform_BigData	Complex	From 1 to 10. **09 Selected**

valuation

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Platform's CSF's

Based on the business case's (and its CSA) LRP4AI process managed and weighted the most important CSFs that were used.

As shown in Table 4, the result's aim is to prove or justify the RDP4AI's feasibility; and the result permits to move to the next CSA that is EA4AI's integration.

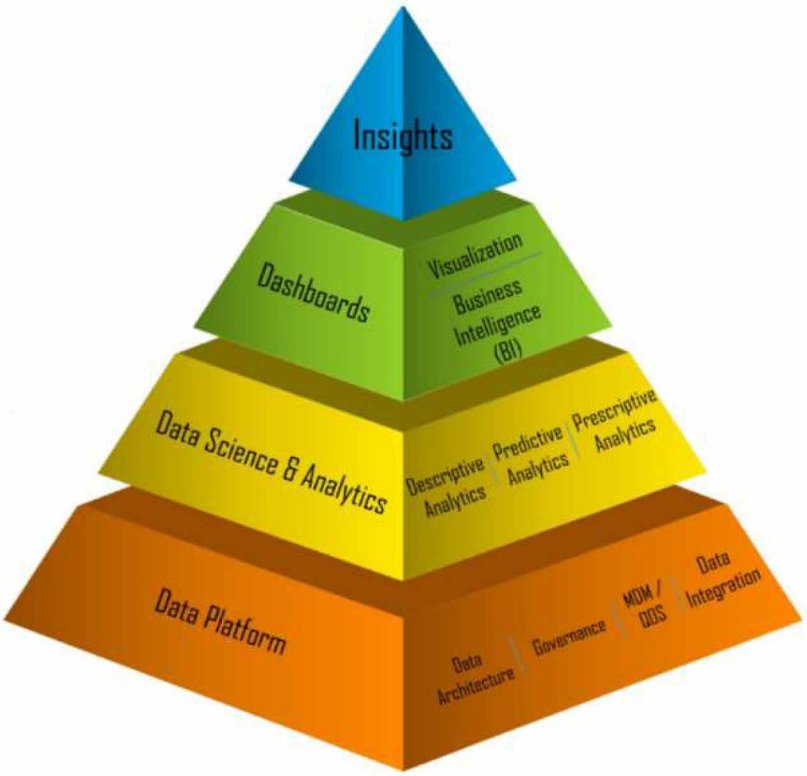
EA4AI'S INTEGRATION

EA4AI Principles and Basics

EA4AI's main principles and basics are:

- DSI4AI may be complex in large *Entities* which with complex risks scenarios.
- EA4AI is often used for strategy purposes and provides an overview of complex *Entity* architectures, showing business entities and relations.
- EA4AI accommodates DMC4AI to enable a data-driven approach, and there is no concept that enables the BGD4AI ro integrate a data-driven EA4AI.
- ADM4AI supports a data-driven *Entity*; through a specific adaption of the ADM4AI.

Figure 14. The layers of data-driven architecture



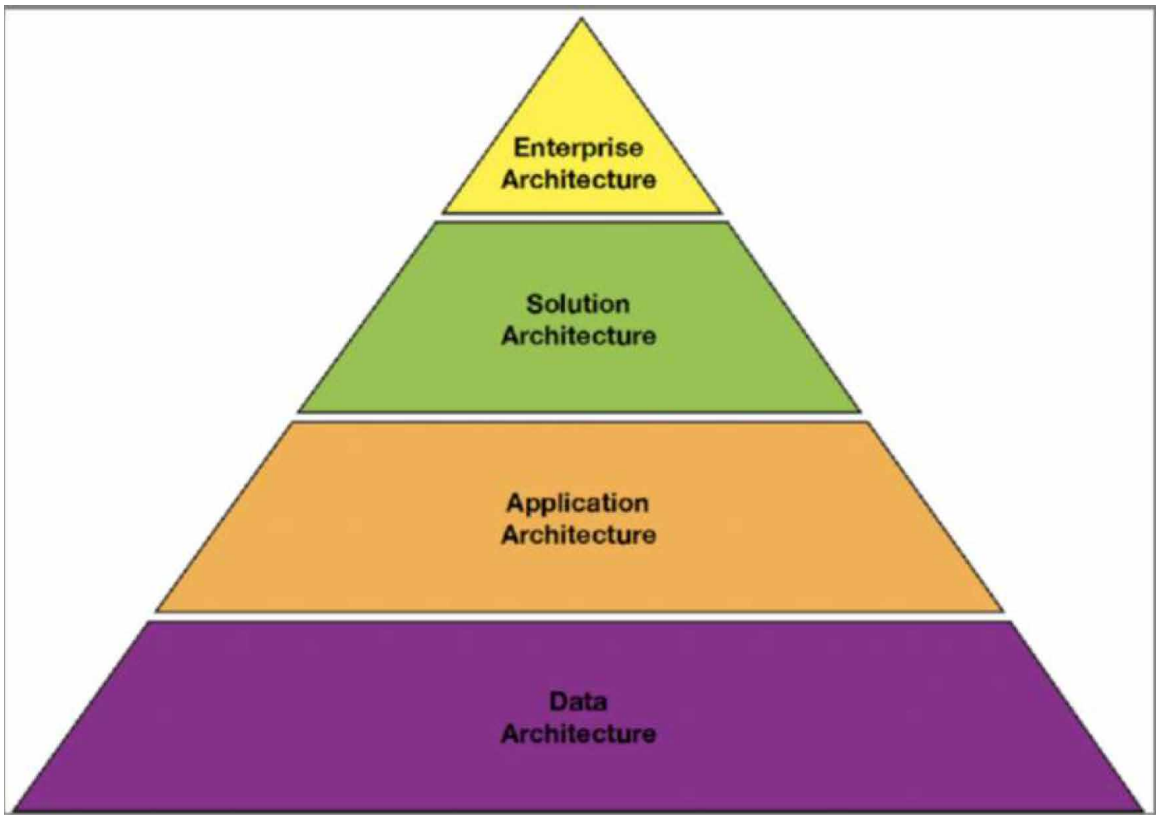
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

EA4AI Layers

EA4AI's main layers characteristics depend on (Sarkar, 2018):

- The lack of reproducibility and reusable artifacts, where the modules are reproducible.
- The lack of collaboration: silos should be removed, even though DSI4AI specialists work in an isolated manner.
- The technical debt: There is a lack of standards and EA4AI concepts for DSI4AI.
- Building reusable assets: Besides the DSI4AI, Non-Functional Requirements are important and includes, scalability, maintainability, availability and other.
- DSI4AI needs to implement an EA4AI approach, to define solutions, application and data models. For this goal there is a need to define layered EA4AI as shown in Figure 15.
- EAI4AI based DSI4AI is capable of: 1) Envisioning end-to-end solutions: 2) To improve and transform the *Entity*.
- Structured evolution is essential to DSI4AI that is an innovative field, that needs GAIPs.

Figure 15. Layered architecture hierarchy (Sarkar, 2018)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

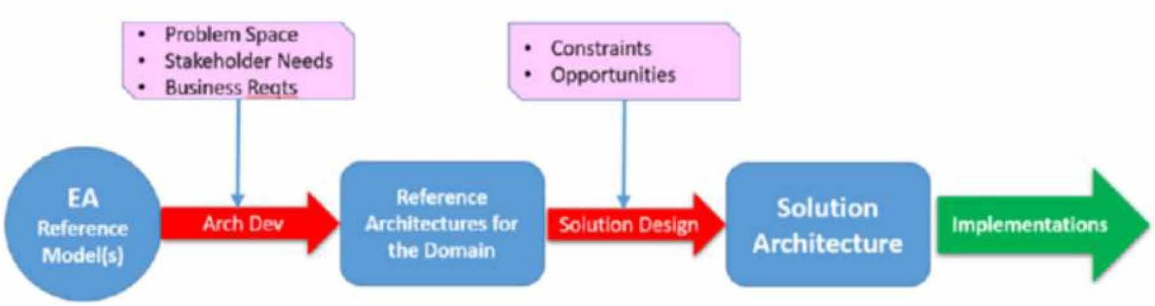
The Role of Data and EA4AI Standards

The *TRADf* uses data technologies and existing data standards that include: 1) RDBMS standards; 2) vertical industry documents' standards; 3) business process standards; 4) data governance standards; 5) business services standards; 6) extensible mark-up language standards; 7) Object Oriented (OO); and 8) resources description standards. Standardization for inter-operability can be achieved by using an ESB and its ETL module (Tamr, 2014). The evolution of standards, like the SOA standards, have enabled DSI4AIs to become more receptive to development and integration with various standards. Standardized *Project*s have to be inter-operable and their focus must be on their: 1) data models; 2) data architectures; 3) modelling concepts; and 4) data monitoring platforms. Regardless of the business domain, executive management understands the immense need for agility and the integration of DSI4AIs. The DSI4AI helps in establishing a unique GAIP.

Application Reference Model

Application Reference Model (ARM) categorizes the system, applications' standards and ICS that support the delivery of service capabilities, allowing various *Entities* to share common solutions. EA4AI's data architect and a *Project's* data analyst, can use the standard ARM, which is not on for analytics, but is also used for EAI4AI. ARM is a used, for mapping business APD to *Entity's* applications. ARM is a key for predictive analytics initiatives, where it defines the interface between the business and the ICS that supports it. GAIP needs to create ARM artefacts to be used by the DSI4AI (Doree, 2015).

Figure 16. The model's implementation



EA4AI's CSF's

Based on the business case's (and its CSA) LRP4AI process managed and weighted the most important CSFs that were used.

As shown in Table 5, the result's aim is to prove or justify the RDP4AI's feasibility; and the result permits to move to the next CSA that is modelling and implementation environments.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. The factors have an average of 9.0

Critical Success Factors	KPIs	Weightings
CSF_EA4AI_Integration_Principles	Feasible	From 1 to 10. **09 Selected**
CSF_EA4AI_Integration_Integration_Strategy	Feasible	From 1 to 10. **09 Selected**
CSF_EA4AI_Integration_Standards	Feasible	From 1 to 10. **09 Selected**
CSF_EA4AI_Integration_EA4AI	Proven	From 1 to 10. **10 Selected**
CSF_EA4AI_Integration_CSF_CSA	Feasible	From 1 to 10. **09 Selected**
CSF_EA4AI_Integration_Layers	Complex	From 1 to 10. **08 Selected**
CSF_EA4AI_Integration_Modelling_Languages	Proven	From 1 to 10. **10 Selected**
CSF_EA4AI_Integration_ARM	Complex	From 1 to 10. **08 Selected**

valuation

MODELLING AND IMPLEMENTATION ENVURONMENTS

The Basic Approach

The basic approach's main characteristics are (OMNI-SCI, 2021):

- DSI4AI needs capable and autonomous specialists because it is a complex environment.
- DSI4AI needs many skills that include statistical analysis and other fields like:
  - Statistics training, including probability, inferential statistics, linear/vector algebra and calculus.
  - Python and/or R, are DS4AI languages;
  - Apache Spark and SQL.
  - Visualization software.
  - MLI4AI, NLP4AI, DLI4AI.

DATA MODELLING LANGUAGES

EA4AI Modelling Language

EA4AI modelling languages like ArchiMate, which is an open and independent modelling language for *Entity* functional, enterprise and data architectures. ArchiMate is supported by many vendors and consulting companies; and it provides instruments to enable *Entity* architects to describe, analyse and visualize the relationships among business (including data) domains in an clear manner (The Open Group, 2013b).

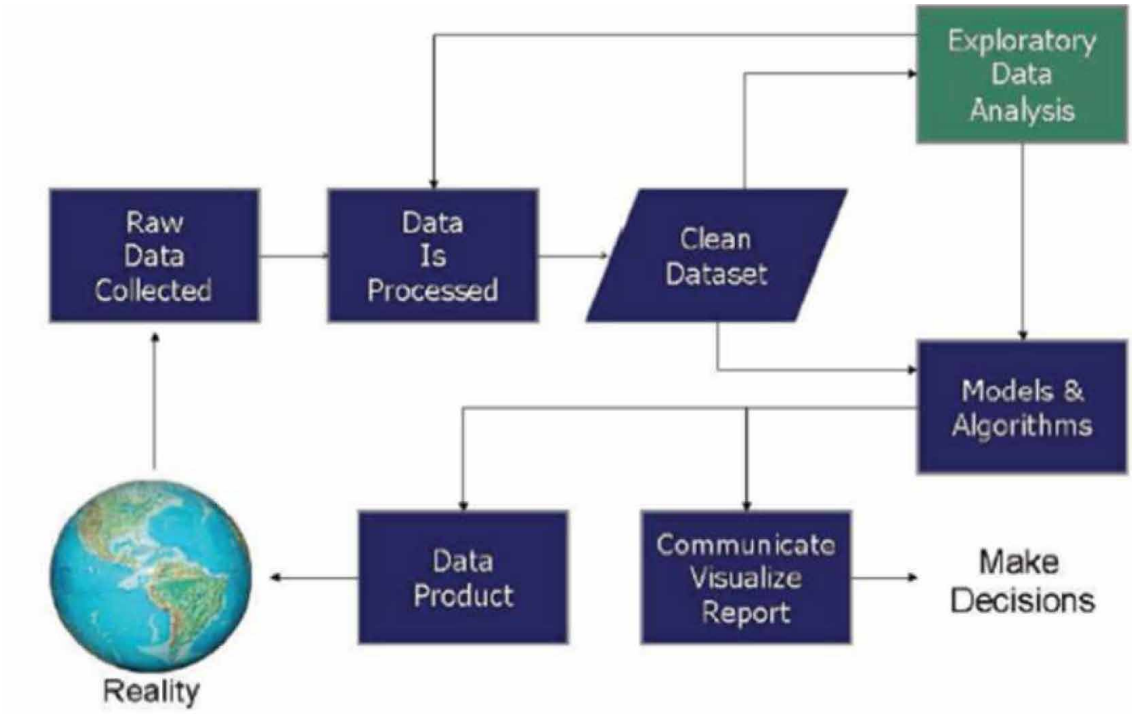
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Unified Modelling Language

Unified Modelling Language (UML) can be used with DSI4AI in the following contexts (Sikander, & Khiyal, 2018):

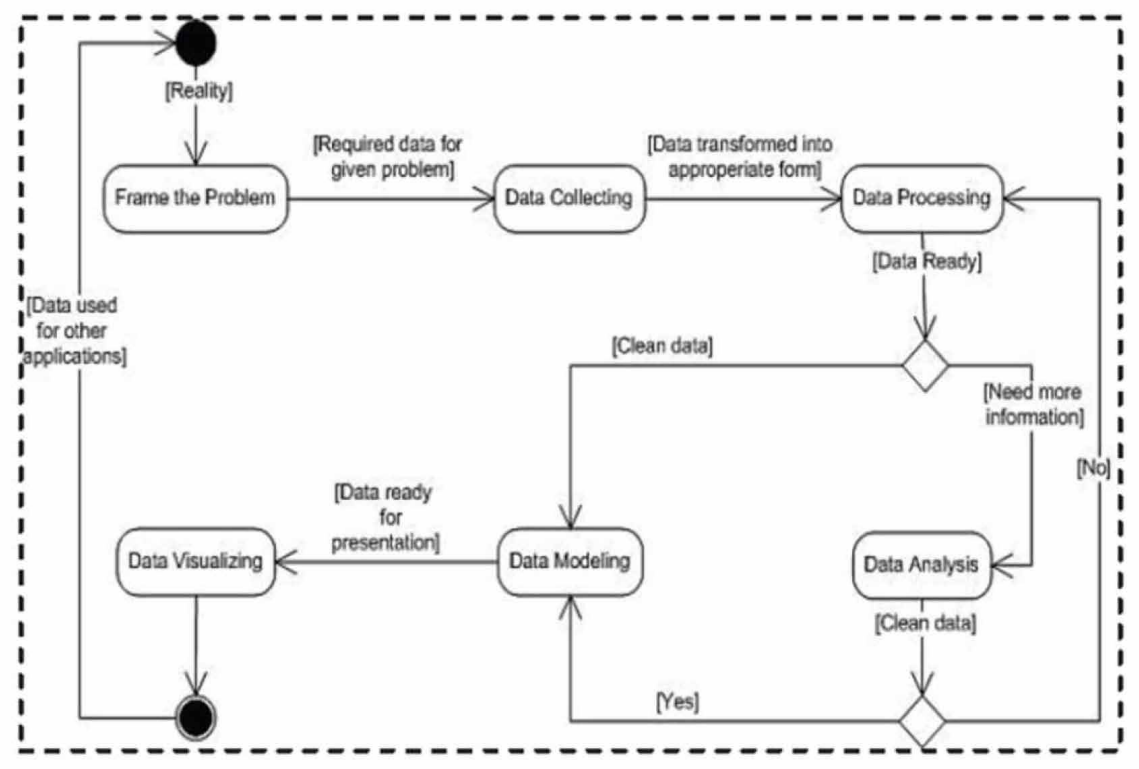
- To model data concepts and implementations; like the Data Flow Diagram (DFD) is an artefact that represents a flow of data through a process or a system (usually an ICS).
- The DFD also provides information on the outputs and inputs of each object (like a table) and the process itself. The DFD has no control flow, there are no Dynamic Rules for AI (DR4AI) and no loops. Specific operations based on the data can be represented by a flowchart.
- A UML Diagram, is essential to illustrate a conceptual model to a precise *Project* problem with the component or class diagrams. And to present the used algorithm, using the sequence or activity diagrams.
- In the context of the process of data analysis, which integrates three models as shown in Figure 17, it is the process of inspecting, cleaning, converting and modelling of data, in order to obtain particular results. Raw data is collected from various sources and transformed into usable information streams for DMS4AI and DSI4AI processing. There are different phases in the process of DMS4AI and DSI4AI processing; these processes are iterative. The UML state diagram for DSI4AI process is shown in Figure 18.

Figure 17. The data analysis process flowchart (Sikander, & Khiyal, 2018)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. The data analysis process flowchart, using the state diagram (Sikander, & Khiyal, 2018)



Entity Relationship Model

An Entity Relationship Model (ERM) describes interrelated objects (tables) of interest in an application domain. A basic ERM is composed of object types and specifies relationships that can exist between objects. In DSI4AI, an ERM is used to represent objects for *Project* needs to remember in order to perform business operations. The ERM is an abstract data model, which defines a data or information structure which can be implemented in a DBMS, commonly a RDBMS.

Data Modelling Diagrams

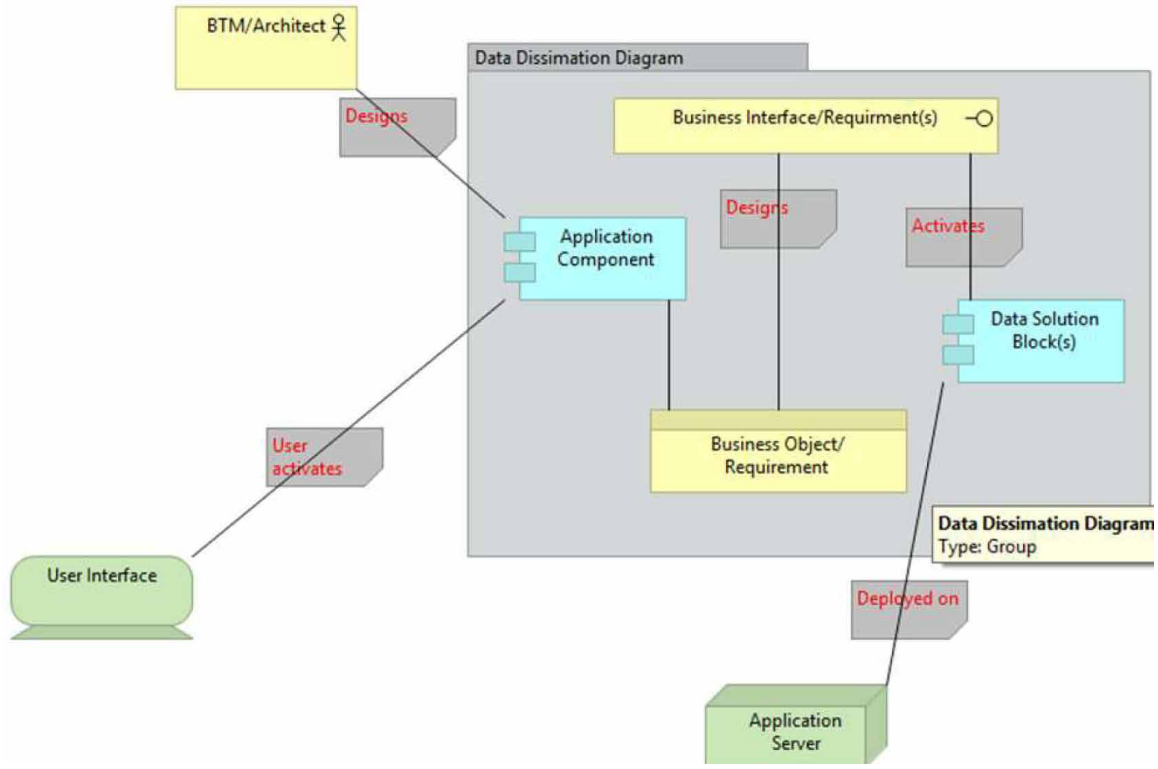
A data model is needed to describe the structure and relationships among a set of data managed in the *Project* and stored in a data management system. The most common is the relational model that is often represented in entity-relationship diagrams or class diagrams; these diagrams basically show data entities and their relationships. The DSI4AI for a *Project* can be seen as a data architectural view (Merson, 2009); where the modelling process incorporates various conceptual views.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Data Dissemination View

The purposed data dissemination diagram shows the relationship within the DSI4AI: 1) data entities; 2) business data services; and 3) application components/services. This diagram as shown in Figure 19, presents the logical data entities to be physically managed/accesses by application components or services. That permits a flexible architecture and modelling of the data sources. This diagram includes data services that abstract various data sources (TOGAF Modelling, 2015a).

Figure 19.



Data Migration Diagram

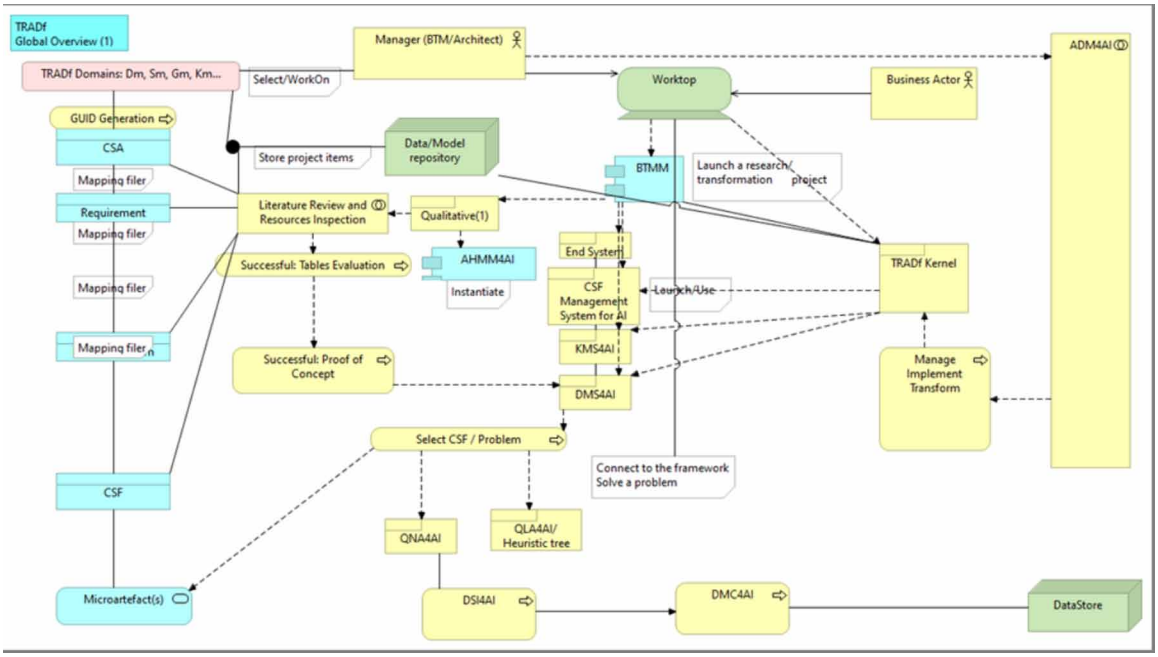
DSI4AI's data migration diagram presents the flow of data between the business systems' various data sources; it serves as a view for the *Project*'s data auditing and traceability map. The data migration diagram is enhanced through *Project*'s iterations (TOGAF Modelling, 2015d). A mapping table was developed to show the relationships between various data points (source, intermediate or destination).

Data Security Diagram

Design of data architecture and modelling are the most important activities in the *Project* and the transformed business data are a very important asset for the *Entity*; that implies that the transformed business

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 20. The experiment's data dissemination view



data security is an ultimate goal. The PoC's data security diagram presents the *Project*'s actors' roles for transformed business data. A matrix is used to present the mapping relation between the role and the transformed data entity. The data security diagram includes all external accesses which may have access to the business system's data (TOGAF Modelling, 2015e).

Functional Development

NLP4AI

TRADf's internal NLP4AI tool can be used for various application domains and in general for hard systems' thinking. The author recommends the use of an interpretable scripting for building a GAIP (Moore, 2014). The AHMM4AI based process is domain-driven and is founded on *TRADf* that in turn is based on a NLP4AI to manage heuristics/rules, *Entity/*enterprise architecture, QLA4AI and QNA4AI microartefacts (Simonin, Bertin, Traon, Jezequel & Crespi, 2010). The complexity lies in how to integrate the AHMM4AI and its programming NLP4AI in GAIP. The main characteristics and facts related to functional languages are (Clancy, 2019):

- Functional programming reaches for efficiency in finance.
- Financial institutions are adopting functional programming or NLP4AI as an alternative to the dominant imperative approach.
- NLP4AI results with less bugs as it is stricter and easier to check and to test.
- It has in-built concurrency constructs, and are optimal for distributed ledgers. Distributed ledgers may find a wide application in the derivatives industry.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- It lacks skilled developers, as well as its memory and speed drawbacks.
- They will become more relevant with the adoption of the Distributed Ledger Technology (DLT) in finance.
- NLP4AI make testing easier because of immutability.
- It will continue to spread, as *Entities* become aware of its advantages.

R vs Python for DSI4AI

DSI4AI languages are used for data transformation, data cleaning, and data visualization. There is also a need to detect outliers, identify relationships between variables, and construct complete interpretive models for DSI4AI. Data preparation and statistical analysis tools like R and Python can be used. R is a language for statistics, data analysis and graphical models.

R

R is particularly suited to data analysis tasks on individual servers that demand standalone analysis or. It's optimal for PoCs and visualization, but it is slow and has a difficult syntax.

Python

Python is a more readable language and is for general use for statistical analysis demands.

F#

F# (F sharp) is a functional, strongly typed, multi-paradigm programming language that incorporates functional, imperative and OO paradigms. F# is a part of the .NET platform.

SCALA

SCALA is a functional and OO programming language it is strong static typed system. Designed to be concise.

Classical Development Tools

Classical development environments, like JEE, C#, C++, Perl, EXCEL/VBA and other, can be used for DSI4AI oriented development.

Environments CSFs

Based on the business case's (and its CSA) LRP4AI process managed and weighted the most important CSFs that were used.

As shown in Table 6, the result's aim is to prove or justify the RDP4AI's feasibility; and the result permits to move to the next CSA that is the GAIP's components integration

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. The factors have an average of 9.40

Critical Success Factors	KPIs	Weightings
CSF_Modelling_Implementation_DSI4AI_Interface	Proven	From 1 to 10. **10 Selected**
CSF_Modelling_Implementation_Modelling_Languages	Proven	From 1 to 10. **10 Selected**
CSF_Modelling_Implementation_Diagrams	Feasible	From 1 to 10. **09 Selected**
CSF_Modelling_Implementation_Functional_Dev	Feasible	From 1 to 10. **09 Selected**
CSF_Modelling_Implementation_CSF_CSA	Proven	From 1 to 10. **10 Selected**
CSF_Modelling_Implementation_ADM4AI_Integration	Proven	From 1 to 10. **10 Selected**
CSF_Modelling_Implementation_APD_Interfacing	Complex	From 1 to 10. **08 Selected**

valuation

Table is wider than 6.25 inches

GAIP'S MAIN COMPONENTS INTEGRATION

GAIP's Construct

The GAIP's construct has the following characteristics and presumptions (OMNI-SCI, 2021):

- AI should simulate the human brain, by using the ICS, which includes learning, logical reasoning and auto-correction.
- The GAIP supports the *Entity's* learning, auto-correct and reasoning processes and the capacity to draw inferences in an independent manner.
- The GAIP is generic and has a holistic approach, because it can handle many types of activities, which demands reasoning, judgment and thought.
- The GAIP uses specialized methods to handle specific tasks.
- GAIP uses neural networks and it needs an extraordinary amount of data.
- For a summary: 1) DSI4AI is the field; 2) GAIP is the goal; and 3) MLI4AI is the process.

MLI4AI's Integration

The MLI4AI's Integration process has the following characteristics (OMNI-SCI, 2021):

- GAIP based DSI4AI and MLI4AI work together. Where MLI4AI feeds the ICS with huge amounts of data to support DMS4AI's processes.
- The GAIP based DMS4AI uses human interaction to feed the DMS4AI with massive quantities of data, to be able to solve *Project* problems.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Humans feed the DMS4AI with various types of data, so it can learn from data associated with knowledge items/features.
- DSI4AI includes statistics that with ICS technologies can handle massive amounts of data. The MLI4AI is part of the DSI4AI, and it allows the DMS4AI to learn.
- MLI4AI uses algorithms to support the DMS4AI, but those algorithms rely on source data. The DMS4AI uses source data as a learning sets, so it can improve its algorithms. It tunes the parameters of the algorithms by using statistical techniques, including naive Bayes regression and supervised clustering.
- The DMS4AI can train another DMS4AI to detect data structures, by using unsupervised clustering, to optimize the classification algorithm; human support is needed to control the classification of structures that the DMS4AI identifies.
- The scope of DSI4AI goes beyond the MLI4AI, by encompassing data that is not generated by any mechanical process; like in the cases like, survey's data, data from clinical trials or any other kind of data.
- DSI4AI deploys data to train DMS4AIs and is not limited to statistical data processing. The DSI4AI includes automating MLI4AI's data-driven decisions.
- When a DSI4AI specialist creates the insights from data, a DMS4AI learns from those insights. A DMS4AI can build its own insights using existing algorithmic structure, where the starting point relies on some kind of used structured data.

BGD4AI's Integration

The Integration of BGD4AI has the following characteristics (OMNI-SCI, 2021):

- Uses data streams from various sources, like, online purchases, multimedia forms, instruments, financial logs, sensors, text files and others. Data can be: 1) Unstructured; 2) Semi-structured; and 3) Structured.
- Unstructured data, includes data from: blogs, digital audio/video feeds, digital images, emails, mobile devices, sensors, social networks and tweets, web pages and online sources.
- Semi-structured data comes from: system log storages, XML files and text files.
- Structured data, is the data that has been already processed by: OLTP, RDBMS, transaction data processes and other formats' processing.
- BGD4AI's integration is needed to support the processing of tremendous amounts of data from various sources; which is impossible to process with BI or DA4AI tools. BGD4AI based DSI4AI offers *Entities* with advanced, complex algorithms for analysing, cleansing, processing and extracting meaningful insights.
- BGD4AI based DSI4AI is a scientifical research process that uses AHMM4AI and the ICS to process massive data.
- The DSI4AI, includes cross-functional tasks like, data cleansing, intelligent data capture techniques and data mining and NLP4AI support.
- The DSI4AI is capable of capturing, maintaining and preparing BGD4AI for complex and intelligent analysis.
- BGD4AI supports the DSI4AI to deliver velocity, variety and volume (or the 3Vs); that supports the techniques for analysing huge volumes of data.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

THE USE OF STATISTICS

Statistics Methods' Integration

The Integration of statistics methods includes (OMNI-SCI, 2021):

- DSI4AI is a broad, cross functional domain that is used in applied business management, computer science, economics, mathematics, programming and software engineering; which are supported by statistics.
- The DSI4AI requires the collection, processing, management, analysis and visualization of large quantities of data; where DSI4AI specialists use tools from various APDs, including statistics, to achieve those activities.
- There is a close relation between DSI4AI and BGD4AI, where most BGD4AI's data are unstructured and include non-numeric data. Therefore, the task of processing data as a DSI4AI specialist involves removing noise and extracting useful data sets.
- Statistical tasks need specific design and implementation strategies to support: 1) acquisition; 2) architecture; 3) analysis; and 4) archiving processes. These *4As* of DSI4AI are unique to the APD.
- Statistics is a broad field demanding APD expertise and it focuses on the study of numerical and categorical data.
- Statistical theory and methods support the DSI4AI in: 1) gathering data in a performant way; 2) analyse and interpret them for a particular use; and 3) make conclusions to solve particular problems.
- Statistical methods ensure that DSI4AI processes explore and describe data while summarizing them.
- Statistical protocols are essential to deliver accurate prediction and insightful inferences.

Data Mining Technique for AI

The Data Mining Techniques for AI (DMT4AI) has the following characteristics (OMNI-SCI, 2021):

- It is used in business to support the DSI4AI, to render data more usable for a specific business objective.
- It deals with structured data, when exploring huge amounts of unprocessed data.

Deep Learning Integration for AI

The DLI4AI has the following characteristics (OMNI-SCI, 2021):

- It is a function that mimics how the human brain works as it processes data and generates patterns to be used by the DMS4AI.
- It is a type of MLI4AI, that focuses on deep neural networks that can master unstructured or unlabelled data, without any type of human assistance.
- It uses hierarchical artificial neural networks to engage in the MLI4AI process.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Traditional DA4AI uses data in a linear manner, whereas the DLI4AI uses a hierarchy of functions which enables a nonlinear approach to problems' solving.

DSI4AI for Business

The DSI4AI for business has the following characteristics (OMNI-SCI, 2021):

- Supports analytics that is used in an APD.
- DSI4AI supports businesses to understand better the specific needs of customers by using the existing data.
- DSI4AI processes can deliver training models for searching and delivering recommendation.

Business Analytics Integration for AI

The Business Analytics Integration for AI (BAI4AI) has the following characteristics (OMNI-SCI, 2021):

- The DSI4AI and BAI4AI focus on solving business problems and both involve: 1) Data collection; 2) Modelling data; and 3) Gleaning insights from the data. The main difference is that BAI4AI is specific to business-related problems, like, profit and costs.
- In contrast, DSI4AI methods explore how a wide range of CSFs might affect the *Entity's* business. DSI4AI combines data with ICS and algorithms' implementation to solve various types of problems.
- BAI4AI is a specialized field, analysing data from the business using traditional statistical theory to generate insights and to deliver business solutions.

Business Intelligence Integration for AI

The BII4AI has the following characteristics (OMNI-SCI, 2021):

- It is a subset of DAI4AI and it analyses existing data for insights to discover business trends.
- It gathers data from internal and external sources to: 1) prepare processes for a specific use; and 2) creates dashboards to resolve business problems.
- BII4AI tools can usually evaluate specific events which might affect an *Entity* in the future.

DSI4AI in Finance

The DSI4AI in finance has the following characteristics (OMNI-SCI, 2021):

- It is a powerful tool to combat fraud, giving financial institutions to recognize problematic patterns in data.
- It helps to reduce non-performing assets and to reveal downward trends.
- DSI4AI creates a model that can perform predictive analytics on various topics.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

DSI4AI Use in Transformation Initiatives

The DSI4AI in transformation initiatives has the following characteristics (OMNI-SCI, 2021):

- An *Entity* can use DSI4AI to shape policies to meet the needs of their constituents, by collecting data with the census, surveys and tools.
- Correlative or descriptive analysis of used data sets is used to support the DMS4AI.

GAIP's CSF's

Based on the LRP4AI process the most important CSFs that were used.

Table 7. The factors have an average of 8.50

Critical Success Factors	KPIs	Weightings
CSF_DSI4AI_GAIP_Construct	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_GAIP_ML4AI_DLI4AI	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_GAIP_Analytics_Statistics	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_GAIP_BGD4AI	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_GAIP_DMT4AI	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_GAIP_Business	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_GAIP_Finance	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_GAIP_Transformation	Complex	From 1 to 10. **08 Selected**

valuation

As shown in Table 7, the result's aim is to prove or justify GAIP's feasibility; and the result permits to move to the next CSA that is the assembling the DSI4AI component.

ASSEMBLING DSI4AI'S COMPONENTS

DSI4AI components enable *Entities* to process huge amounts of structured and unstructured data from BG4AI to detect patterns. This enables them to increase efficiencies, manage costs, identify new market opportunities and boost their market advantage (OMNI-SCI (2021).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

THE ANALYSIS PROCESSES

The Process

DSI4AI's Application Programming Interface (API) can be used to avoid common pitfalls, because of expensive, complex and monolithic nature, the barrier to integrate DSI4AI has been simplified with the introduction of simpler and distributed platform becomes viable (Gartner, 2020).

People

In order to motivate stakeholders, it is imperative that there is a strong relationship between the DSI4AI and the *Entity's* business results. This is complex to achieve, because of ICS' silo structure (Gartner, 2020).

Technology

To assess impacts related to *Projects*, firstly is important to assess DSI4AI's solutions for a given problem and whether the *Entity* is ready to be apply the solution. It is recommended to start with a PoC and assess an end solution. Secondly, as it is imperative not to start with ICS/technology integration, and where EA4AI needs to inspect the existing *Entity's* processes significantly (Gartner, 2020).

Data-Driven Business Transformation Projects

Data-driven *Projects* (or simply *DDProject*) has the following characteristics (IBM, 2019):

- DMS4AI based on data is daily business and there is a need to adopt new GAIP oriented technologies and methodologies that facilitate data-centric DMS4AI; that is the first important step for a *Project*. Where GAIP is the skeleton of MLI4AI and DLI4AI.
- DSI4AI and GAIP are synergistic, where GAIP is the skeleton of the DSI4AI, which improves data querying accuracy and performance and helps to optimize *Project* resources. DSI4AI platforms are evolving to support GAIP like patterns.
- Data platforms and GAIP analytics references from the LRP4AI, reveal the extent to which *Entities* use GAIP and MLI4AI as critical goals of their data platform and analytics initiatives. Two-thirds of *Entities* confirm that a GAIP like approach and MLI4AI are important components of their data platform and DA4AI initiatives. The increase upto 88% among the most data-driven *Entities*, where strategic *Project* decisions are data-driven.

Main Components

DS4AI main components are (Guru99, 2021):

- Statistics is the most critical unit of DSI4AI and it is the method or science of collecting and analysing numerical data in large quantities to deliver insights.
- Visualization technique supports the access to huge amounts of data in digestible visuals.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- MLI4AI explores the building and study of algorithms which learn to deliver predictions about unforeseen/future data.
- DLI4AI method is a new MLI4AI based research where the algorithm selects the analysis model to be followed.

DSI4AI Processes

The main DSI4AI processes are (Guru99, 2021):

- The *Discovery* process, involves data acquisition from the *Entity's* sources and supports various types of requests. Where data can be: 1) Log files from webservers; 2) Data gathered from social media; 3) Census datasets; and 4) Data streamed from online sources using APIs
- The *preparation* of data can generate inconsistencies like, missing values, blank columns, incorrect data formats; which needs to be corrected or removed. Correction implies, to process, explore and condition data models before modelling. The cleaner the *Entity's* data are, the more credible are the made predictions.
- The *Model Planning* process, determines the method and technique to define the relation between input variables. Planning for a model is done applying different statistical formulas and visualization environments.
- The *Model Building* process, prepares datasets for training and testing are distributed. Modelling techniques like association, classification and clustering are applied to the training data sets. Once the model prepared, it is tested against the *testing* datasets.
- The *Operationalization process*, delivers the final baseline models with reports, code and technical documents. The models are deployed in production environments after testing has terminated.
- *Communicating results*, to all stakeholders, shows the results and if they are successful or not.

DS4AI NEEDED SKILLS

DS4AI Roles

DSI4AI's main skill/specialist roles are:

- Data Scientist.
- Data Engineer.
- Data Analyst.
- Statistician.
- Data Architect.
- Data Admin.
- Business Analyst.
- Data/Analytics Manager.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Certifications

Today, there are various certifications, like, The Open Group Open Certified Data Scientist (Open CDS) certification, which is the industry's most comprehensive certification available for Data Scientists. Open CDS is an independent global certification for qualifying the skills, knowledge, and experience of Data Scientists (The Open Group, 2021).

DSI4AI's CSF's

Based on the LRP4AI process the most important CSFs that were used.

Table 8. The factors have an average of 8.70

Critical Success Factors	KPIs	Weightings
CSF_DSI4AI_Assembling_Analysis_Processes	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_Assembling_DDProject	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_Assembling_Integration_Components	Complex	From 1 to 10. **08 Selected**
CSF_DSI4AI_Assembling_EA4AI_ADM4AI	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_Assembling_CSF_CSA	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_Assembling_Processes	Feasible	From 1 to 10. **09 Selected**
CSF_DSI4AI_Assembling_Skills	Proven	From 1 to 10. **10 Selected**
CSF_DSI4AI_Assembling_Domains	Feasible	From 1 to 10. **09 Selected**

valuation

As shown in Table 8, the result's aim is to prove or justify the RDP4AI's feasibility.

THE PROOF OF CONCEPT

Basics

TRADf's Tm empirical model was built to prove this RDP4AI's hypothesis by using a concrete PoC, which has been developed using the Microsoft Visual Studio 2019 and Python. The PoC contains *TRADf*'s major components needed for the BIG4AI based DSI4AI's processing and was primarily tested its mixed reasoning engine, which is based on the heuristics model. This PoC serves to confirm the research's RQ. The reasoning engine has a defined goal function, which calculates the best solution

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

for the encountered DSI4AI problems or requests. The PoC's results are presented in the form of a set of technical and managerial recommendations.

CSFs, Rules and Constraints Setup

The DSI4AI's process execution starts with the use of the imputed data collection in *TRADf's* massive data storage and then these data are filtered using the selected set of CSFs. The execution of the qualitative part follows; initially inputted *TRADf's* data collection is considered to be the root or initial node that helps in the establishment of the basic set of possible problems and to be eventually enhanced with the adopted DSI4AI's solution(s). The DMS4AI's tree reasoning goal is to select the optimal solutions and to implement them in *Project* successfully.

The Tree and Resources

GAIP's decision tree's collection of nodes contains the following resources, which are based on the unbundling of the system: 1) The executed actions; 2) The constraints; 3) The DSI4AI related problems; and 4) The set of possible solutions. Each tree node is a CSF suggestion and it is linked to a concrete data state, which in turn contains an aggregate of a resource linked with a 1:1 mapping link.

Possible Solutions

The selected CSFs were fed in the DMS4AI's heuristics engine in order to reveal the optimal DSI4AI prerequisites for a selected *Project* problem or request. The CSFs were configured and weighted; afterwards they were processes in order to deliver a set of possible solutions. The DMS4AI starts with the initial set of selected CSFs that correspond to a specific problem or request; then the grounded hyper-heuristics processing is launched to find a set of possible solution(s) in the form of possible improvements or suggestions of needed actions (Jaszkiewicz, & Sowiński, 1999). The author's aim is to convert their relevant research outcomes into a set of managerially useful recommendations for DSI4AI's integration; and *TRADf's*hyper-heuristics tree processing model template that is suitable for a wide class of problem instances (Vella, Corne, & Murphy, 2009).

The PoC Processing

The PoC uses an internal set of CSFs' that are presented in Tables 1 to 8. These CSFs have bindings to specific RDP4AI resources, where the AHMM4AI formalism was designed using an NLP4AI micro-artefacts. In this chapter's tables and the result of the processing of the DMS4AI, as illustrated in Table 9, shows clearly that the DSI4AI is feasible.

The *TRADf* and hence the AHMM4AI's main constraint is that CSAs for simple research components, having an average result below 8.5 will be ignored. In the case of the DSI4AI's implementation an average result below 6.5 will be ignored. As shown in Table 9 the average is 9.15. The BGI4AI based DSI4AI processing model represents the relationships between this research's requirements, NLP4AI generic and microartefacts, unique identifiers and the CSAs. The PoC was achieved using *TRADf's* client interface. From the *TRADf's* client interface, the NLP4AI development setup and editing interface can

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9. The research's outcome

	Critical Success Factors	KPIs	Weightings Ranges	Values			
1	RDP5AI	HighlyFeasible	From 1 to 10.	9,25			
2	AHMM4AI	PossibleClassification	From 1 to 10.	10		EvaPA	
3	DMC4AI	AutomatedExists	From 1 to 10.	9,4			
4	PLATFORM	IntegrationPossible	From 1 to 10.	9			
5	EA4AI	AdvancedStage	From 1 to 10.	9			
6	ENVIRONMENT	Advanced	From 1 to 10.	9,4			
7	GAIP	IntegrationPossible	From 1 to 10.	8,5			
8	ASSEMBLING	IntegrationPossible	From 1 to 10.	8,7		RESULT:	9,15625

be launched. Once the development setup interface is activated the NLP4AI interface can be launched to implement the needed microartefact scripts to process the defined three CSAs. These scripts make up the kernel knowledge system and the DMS4AI set of actions that are processed in the background. The DSI4AI and the DMS4AI that automatically generates actions which make calls mixed (QLA4AI and QNA4AI) modules, that manages the edited NLP4AI script and flow. DSI4AI structures, serve various APDs and in this PoC the functional domain are information analysis and decision making of a business transaction's log; where the data that results from business transactions are logged and used in the data analysis process.

Information Analysis

DSI4AI's PoC took into account the integration of DAI4AI processes that are used by a virtual user who simulates a business analyst. In this PoC a spreadsheet was built to use various data sources that fed the prototyped business transactions system.

Layers

The prototyped business transaction is based on the PoC's class model; The data architecture and modelling PoC's layers are:

- Data architecture using TOGAF to link DSI4AI to various artefacts.
- The client layer that contains the following packages based on: RDBMS, Microsoft Excel, Flat Files Interfaces and a Client web interface.
- The data services layer that contains the data services hub.
- The data management layer that contains the following packages: an entity relational model, an extensible mark-up language transformer and a NoSQL database interface.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- The control-monitoring layer contains a generic logger interface.
- The platform layer that contains the following packages: the *Entity* ESB, an object database connector and an application database connector.

SOLUTIONS AND RECOMMENDATIONS

This chapter's and its PoC's list of the most important technical and managerial recommendations:

- The *Manager*s should be supported with a set of BGI4AI based DSI4AIs tools that can be easily integrated with the defined data-driven EA4AI principles (IBM, 2015a).
- The role of data standards is important; today there are many standards concerning the DSI4AI. The *Manager* must propose a solution on how to integrate these various data standards in the *Project*.
- GAIP's data service bus and *Entity* application integration concept must be defined for the *Project*. That would glue the various data sources of the business environment.
- The meta-model and the corresponding "1:1" mapping approach can be used to manage an agile *Project* process (OASIS, 2014).
- Unbundling through the use of business data services uses the DMS4AI to integrated in the context of a system that applies a holistic approach (Daellenbach, McNickle, & Dye, 2012).
- The data services and dBB4AIs must unify the implementation and usage of data models.
- To define BGI4AI's and DSI4AI's default CSFs.
- Implement an ETL process to access data from various locations to enable their unification and normalization for DSI4AI usage.
- Data transformation models need a meta-model to show the relationships between various types of data models and data sources.
- Define a data normalization and inter-operability paradigm that must be based on various standards.
- Define a business process and a data-driven EA4AI and modelling integration paradigm to enhance the KMS4AI and DMS4AI.
- Develop an agile MVC4AI pattern that is dynamically created for each business transaction. This MVC4AI uses a dBB4AI and its instance
- Define the ADM4AI's integration with the proposed concept where the data building and solution blocks are the basic artefacts that circulate through the ADM4AI.
- Define business architecture integration paradigm that is needed to manage the implementation of the dBB4AIs (The Open Group, 2014a).
- Define conceptual views that can be built to simplify the application of the DSI4AI. Such views can also be the used to simplify the *Entity*'s data-model that links various data sources.
- Create basic class diagrams for the DSI4AI(s) to be the central artefact that is defined in the initial phase and then calibrated in all other phases.
- Define the usage of business transactions' data models in various class diagrams.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

This RDP, which manages BGI4AI and DSI4AI modules, is mainly motivated by high the failure rates and complexity in *Projects*. The main activities are related to data architecture and modelling components for a *Project*. A general data concept is needed to finalize the implementation phase of a *Project* and the integration of BGI4AI. There are several trends shaping the future of DSI4AI: 1) the DSI4AI tasks in the *Project* life-cycle are automated and that increases the *Entity's* ROI: 2) the DSI4AI feasibility should be verified by a PoC; 3) another important element is that, DSI4AI resources become accessible to a large set of end-points; 4) to manage the tension between the right to privacy, the need to regulate and the state of transparency; and 5) DSI4AI has to capacity to integrate MLI4AI algorithms transparent, that makes regulatory oversight possible. To implement an AI and Bigdata based data sciences for *Projects* it is recommended to establish a holistic framework like *TRADf* and an EA methodology like TOGAF. This whole chapter is a pattern on how to build such a system; added to that there are many recommendations.

FUTURE RESEARCH DIRECTIONS

The *TRADf* future research will focus on MLI4AI's internals.

REFERENCES

Camarinha-Matos, L. M. (2012). *Scientific research methodologies and techniques- Unit 2: Scientific method. Unit 2: Scientific methodology. PhD program in electrical and computer engineering*. Uninova.

Clancy, L. (2019). *Functional programming reaches for stardom in finance*. CME Group. Accessed and reviewed in April 2021, https://www.risk.net/risk-management/6395366/functional-programming-reaches-for-stardom-in-finance

Daellenbach, H., McNickle, D., & Dye, Sh. (2012). *Management Science. Decision-making through systems thinking* (2nd ed.). Plagrave Macmillian.

Dantzig, G. (1949). Programming of Interdependent Activities: II Mathematical Models. *The Journal of the Operational Research Society*, *53*(11), 1275–1280.

Doree, J. (2015). *Data Science, Data Architecture-Predictive Analytics Methodology: Building a Data Map for a Greenfield Data Science Initiative*. Datasciencearchitect. Accessed and reviewed in April 2019, https://datasciencearchitect.wordpress.com/tag/togaf/

Fodor, O., Dell'Erba, M., Ricci, F., Spada, A., & Werthner, H. (2002). *Conceptual Normalisation of XML Data for Interoperability in Tourism. Commerce and Tourism Research Laboratory ITC-irst*. University of Trento.

Gartner. (2020). *Three Essentials for Starting and Supporting Master Data Management*. ID G00730039. Gartner Inc. Accessed and reviewed in April 2019, https://www.gartner.com/doc/reprints?id=1-24MIGFGU&ct=201119&st=sb

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Guru99. (2021). *Data Science Tutorial for Beginners: What is, Basics & Process*. Accessed and reviewed in April 2019, https://www.guru99.com/data-science-tutorial.html

IBM. (2015a). *Modelling the Entity data architecture*. IBM.

IBM. (2019). *AI and Data Management-Delivering Data-Driven Business Transformation and Operational Efficiencies*. IBM USA.

IMD. (2015). *IMD business school and Cisco join forces on digital business transformation*. IMD.

Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality & Quantity, 41*(1), 37–54. https://link.springer.com/article/10.1007/s11135-005-5427-1

Jaszkiewicz, J., & Sowiński, R. (1999). The 'Light Beam Search' approach - an overview of methodology and applications. *European Journal of Operational Research*, *113*(2), 300–314. doi:10.1016/S0377-2217(98)00218-5

Kim, K., & Kim, K. (1999). Routing straddle carriers for the loading operation of containers using a beam search algorithm. Elsevier. *Computers & Industrial Engineering*, *36*(1), 109–136. doi:10.1016/S0360-8352(99)00005-4

Lazar, I., Motogna, S., & Parv, B. (2010). Behaviour-Driven Development of Foundational UML Components. Department of Computer Science. Babes-Bolyai University, Cluj-Napoca, Romania. doi:10.1016/j.entcs.2010.07.007

Leitch, R., & Day, Ch. (2000). *Action research and reflective practice: towards a holistic view*. Taylor & Francis. https://www.tandfonline.com/doi/ref/10.1080/09650790000200108?scroll=top

Lockwood, R. (2018). *Introduction The Relational Data Model*. Accessed and reviewed in April 2021, http://www.jakobsens.dk/Nekrologer.htm

Loginovskiy, O. V., Dranko, O. I., & Hollay, A. V. (2018). *Mathematical Models for Decision-Making on Strategic Management of Industrial Enterprise in Conditions of Instability*. In Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications (EMIT 2018), Leipzig, Germany.

Mehra, A., Grundy, J., & Hosking, J. (2005). A generic approach to supporting diagram differencing and merging for collaborative design. In *ASE '05 Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering*. ACM. 10.1145/1101908.1101940

Merson, P. (2009). *Data Model as an Architectural View. Technical Note. CMU/SEI-2009-TN-024*. Research, Technology, and System Solutions.

Modelling, T. O. G. A. F. (2015a). *Data dissemination view*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-dissemination-diagrams-menu.html

Modelling, T. O. G. A. F. (2015c). *Data lifecyle diagram*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-lifecycle-diagrams-menu.html

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Modelling, T. O. G. A. F. (2015d). *Data migration diagram*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-migration-diagrams-menu.html

Modelling, T. O. G. A. F. (2015e). *Data security diagram*. TOGAF Modelling. Accessed and reviewed in April 2021,http://www.TOGAF-modelling.org/models/data-architecture-menu/19-data-security-diagrams.html

Moore, J. (2014). *Java programming with lambda expressions-A mathematical example demonstrates the power of lambdas in Java 8*. Javaworld. Accessed and reviewed in April 2021, https://www.javaworld.com/article/2092260/java-se/java-programming-with-lambda-expressions.html

Nijboer, F., Morin, F., Carmien, S., Koene, R., Leon, E., & Hoffman, U. (2009). Affective brain-computer interfaces: Psychophysiological markers of emotion in healthy persons and in persons with amyotrophic lateral sclerosis. In *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE. 10.1109/ACII.2009.5349479

OASIS. (2009). *OASIS Web Services Reliable Messaging (WS-ReliableMessaging) Version 1.2*. OASIS Standard. Accessed and reviewed in April 2021, http://docs.oasis-open.org/ws-rx/wsrm/200702/wsrm-1.2-spec-os.html

OASIS. (2014). *ISO/IEC and OASIS Collaborate on E-Business Standards-Standards Groups Increase Cross-Participation to Enhance Interoperability*. The OASIS Group. Accessed and reviewed in April 2021, https://www.oasis-open.org/news/pr/isoiec-and-oasis-collaborate-on-e-business-standards

OMNI-SCI. (2021). *Data Science - A Complete Introduction*. OMNI-SCI. Accessed and reviewed in April 2021, https://www.omnisci.com/learn/data-science

Sarkar, D. (2018). *Get Smarter with Data Science — Tackling Real Enterprise Challenges. Take your Data Science Projects from Zero to Production.* Towards Data Science. Accessed and reviewed in April 2021, https://towardsdatascience.com/get-smarter-with-data-science-tackling-real-enterprise-challenges-67ee001f6097

SAS. (2021). *History of Big Data*. SAS. Accessed and reviewed in April 2021, https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Sikander, M., & Khiyal, H. (2018). *Computational Models for Upgrading Traditional Agriculture*. Preston University. Accessed and reviewed in April 2021, https://www.researchgate.net/publication/326080886

Simonin, J., Bertin, E., Traon, Y., Jezequel, J.-M., & Crespi, N. (2010). Business and Information System Alignment: A Formal Solution for Telecom Services. In *2010 Fifth International Conference on Software Engineering Advances*. IEEE. 10.1109/ICSEA.2010.49

Tamr. (2014). *The Evolution of ETL.* Tamr. http://www.tamr.com/evolution-etl/

The Open Group. (2011). *Open Group Standard-TOGAF® Guide, Version 9.1*. The Open Group.

The Open Group. (2013b). *ArchiMate®.* Accessed and reviewed in April 2021, http://www.opengroup.org/subjectareas/*Entity*/archimate

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The Open Group. (2014a). *The Open Group's Architecture Framework-Building blocks-Module 13*. Accessed and reviewed in April 2021, www.open-group.com/TOGAF

The Open Group. (2021). *The Open Group Professional Certification for the Data Scientist Profession (Open CDS)*. The Open Group. Accessed and reviewed in April 2021, https://www.opengroup.org/certifications/certified-data-scientist-open-cds

Trad, A., & Kalpić, D. (2020a). *Using Applied Mathematical Models for Business Transformation. Author Book*. IGI-Global. doi:10.4018/978-1-7998-1009-4

Trujillo, J., & Luj'an-Mora, S. (2003). *A UML Based Approach for Modelling ETL Processes in Data Warehouses*. Dept. de Lenguajes y Sistemas Inform'aticos. Universidad de Alicante.

Vella, A., Corne, D., & Murphy, C. (2009). *Hyper-heuristic decision tree induction*. Sch. of MACS, Heriot-Watt Univ., Edinburgh, UK. Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress.

Veneberg, R. (2014). *Combining enterprise architecture and operational data-to better support decision-making. University of Twente*. School of Management and Governance.

ADDITIONAL READING

Farhoomand, A. (2004). *Managing Business Engineering transformation*. Palgrave Macmillan.

KEY TERMS AND DEFINITIONS

Data Modelling: Is the process of developing data models for the business transformation process. **CSFs:** can be used to manage the statuses and gaps in various project plans and gives the Projects the capacity to proactively and automatically recognize erroneous aBB4AIs and to just-in-time reschedule the Project plan(s).

*TRADf***:** Is this research's framework.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 23
The Effect of Population, Public Investments, and Economic Growth on Regional Renewable Energy Consumption in Turkey: Dynamic Panel Data Analysis

Özlem Karadağ Albayrak
https://orcid.org/0000-0003-0832-0490
Kafkas University, Turkey

ABSTRACT

Making the use of renewable energy sources widespread is of paramount importance for Turkey as for all countries. In this regard, the determiners of renewable energy consumption have been investigated. The effect of determining or factors affecting the use of renewable energy sources on a regional scale to Turkey were examined with different qualitative and quantitative research techniques. In this study, the factors of economic growth, public investments, and population are analyzed by considering regional differences on the consumption of renewable energy resources. The effect of regional economic growth, regional public investments, and regional population on the amount of regional renewable energy consumption were investigated by using panel data of 26 statistical regions of Level-2 classification in the period between 2010-2018 in Turkey. The results obtained by the dynamic panel data analysis concluded that economic growth and public investments at the regional level increased renewable energy consumption while the population growth decreased.

INTRODUCTION

Renewable energy sources are the safest alternatives to fossil fuels produced from oil, natural gas and coal. The reasons for searching alternatives to fossil fuels may be stated as their being generally exhausted and their damage to the environment. Therefore, renewable energy sources as an alternative to

DOI: 10.4018/978-1-7998-6985-6.ch023

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

fossil fuels in energy production are of paramount importance for all countries with different levels of development. The use of fossil fuels in large quantities in energy generation since the Industrial Revolution has supported the development of social productivity to a great extent but has caused important problems such as environmental pollution (Wang Q et. al., 2020). Energy use increases the economic growth and development of a country but results in the problems, such as global warming and carbon emission (Khan H. et al., 2020). Fossil sources such as oil, natural gas and coal, used widely in energy production, will run out of. This problem poses a threat to the sustainability of energy supply. The renewable energy consumption influences sustainable development positively and in a statistically significant manner (Güney, 2020).

Environmental agreements, to which many countries are a party, such as the Kyoto protocol, have also encouraged the search for alternatives to fossil fuels (Sasmaz et. al, 2020). In countries where the share of non-renewable energy sources in the overall energy mix is higher, the increase in energy consumption results in an increase in carbon emissions (Pata, 2018). Accordingly, large industrialized countries play a key role in the solution to this problem as they use most of the fossil fuels produced and technological development (Byrne et. al., 2007).

Increasing demand for energy consumption, strong commitments to remove environmental concerns, technology development and political support will enable the renewable energy an increasingly important role in the global energy mix in the long term (Anton and Nucu, 2020). The renewable energy sources, such as solar, wind, hydraulic, geothermal, biomass, wave, hydrogen either do not cause or result in very little emissions causing global warming. In this respect, encouraging the use of renewable energy sources is of utmost importance for providing a sustainable economic growth in environmental respects (Le et al., 2020).

While the energy supply in the world was obtained from oil, coal and natural gas respectively between 1990-2018, the supply provided from the renewable energy sources was least (International Energy Agency, 2021). As in many other countries across the world, Turkey also has limited fossil fuel reserves. Being dependent on outside financial sources regarding fossil fuels results in an economic burden for countries. In addition, any political tension between countries poses a threat to energy supply. For this reason, the ability to create their own resources using renewable energy resources provides social, economic and environmental benefits for countries. For example, the benefits of employment generation cover more than the income obtained from employment, produced goods and services due to "multiplier effect" (Le et al., 2020).

In 2019, Turkey provided approximately 45.2% of its energy production from renewable energy sources. A large part of this ratio has been generated by hydraulic sources with dams, which are not defined in YEKDEM legislation. Therefore, most of energy demand in Turkey is supplied by fossil sources. In general, policy-makers conduct research in many different fields for the increase in renewable energy sources. In this regard, renewable-sourced electricity generators are provided with a guarantee of purchase at certain prices, and there are different supports for the domestic spare parts used. Moreover, Turkey is also a party to environmental agreements such as the Kyoto protocol. Local administrations provide support in this respect by participating in various organizations. In Turkey, İzmir, Konya, Gaziantep as well as Seferihisar, Karşıyaka, Şişli, Kartal, Tepebaşı, Çankaya, Seydişehir, Fındıklı district municipalities are members of the International Council of Local Environmental Initiatives (ICLEI), which is the most common network of the local climate protection cities campaign in the world.

In this study, a regional evaluation was carried out regarding Turkey. All data were obtained in regional level. The goal of the study is to examine the effects of regional gross national product, population and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

public investments on renewable energy consumption, which is an indicator of regional development level with data for the period of 2010-2018 for 26 regions on a regional scale. The need for diversity and localization regarding energy sources is increasingly widely accepted in the world for the following years (Wang B. et. al, 2020).

This study supports the use of dynamic models according to static models while defining regional management behavior (Dahlberg and Johansson, 2000). In this study, the results provided by estimators of Pooled OLS, Fixed Effect, Difference GMM and System GMM are presented all together.

The most significant contribution of this study is to examine the effect of related variables on renewable energy consumption especially in Turkey with different regional development levels. As far as the search researchers have made in the related literature, this study is the first carried out on the regional scale in Turkey. In addition, this study aims to increase information about the role of regional economic growth, public investment and population in encouraging the implementation of renewable sources in energy sector. The conclusions obtained in this study may contribute to the decisions to be taken by policymakers to increase renewable energy consumption, especially in the goals for regional development. Turkey is a country with more regional developmental differences.

Regions such as Istanbul, Ankara and Izmir included in the Level 2 classification have the highest economic growth shares and there is one-way migration movement from other regions towards these regions. Renewable energy resources are the only alternative to eliminate environmental concerns, and they are a very important sector to close the difference between regional development levels economically and socially. This study focuses on regional scale assessment in order to contribute to the sustainable development of the regions. determinants that influence the consumption of renewable energy sources on a regional scale aims to correct this deficiency in the literature and in this context it is very important for Turkey.

Renewable Energy Sources

The renewable energy sources are defined as follows in The Regulation on Certification and Support of Renewable Energy Resources (YEKDEM): “Wind, solar, geothermal, biomass, gas generated from biomass (including landfill gas), wave, current energy, and energy generation production facilities by tidal and canal or river type or reservoir area below fifteen square kilometres or suitable for the establishment of a hydroelectric with pumped storage”. Based on this definition, consumption of energy produced by hydraulic power plants not included in this classification was not subjected in the study.

All renewable energy sources are not used in Turkey. Hydraulics, solar, wind, geothermal, biomass-sourced energy is produced in Turkey. The solar and wind energy generation is carried out by two methods, licensed and unlicensed, and contests are held for licensed generation. There is a quite high potential regarding these sources in Turkey. In 2017, Turkey became the country which increased its installed capacity of solar energy most. It is also ranked among the top five countries in the world regarding geothermal energy generation.

BACKGROUND

Many researchers have conducted investigations on the different factors effective on the amount of renewable energy consumption. In this respect, economic growth, population, public investments, greenhouse

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

gas emissions, financial development, energy prices, trade, foreign direct investment, human capital Index may be given as examples.

Le et al. (2020) analyzed the effect of financial development on renewable energy consumption by means of 2005-2014 panel data of 55 countries by using Two-System generalized moment method (GMM). Omri et al. (2015) investigated the relationship among consumption amount of renewable energy resources with the variables of per capita CO_2 emissions, real oil price, per capita GDP and per capita trade openness by static and dynamic panel data analysis techniques. Padhan et al. (2020) examined the effect of GDP, oil prices and carbon emissions per capita on renewable energy consumption in 30 OECD countries (1970-2015). By utilizing fully modified ordinary least squares (FMOLS) and pooled ordinary least squares (POLS) techniques, Amoah et. al. (2020) analyzed the role of economic welfare and economic freedom as driving forces of renewable energy by means of the panel data of 32 African countries between 1996-2017. Moreover, Assi et al. (2021) analyzed the effect of the variables, such as financial development, environmental pollution, innovation, economic freedom and GDP per capita, specified to be five main factors affecting the renewable energy consumption, by using the panel data belonging to the period of 1998-2018. By using 1980-2018 panel data from 190 countries, Zhao et al. (2020) analyzed the effect of financial development, per capita income and trade openness factors on renewable and non-renewable energy consumption in China by 1980-2015 time series. Khan H. Et al (2020) analyzed the effect of economic growth, foreign direct investment (FDI) and impact of CO_2 emissions on the renewable energy consumption by means of two-step difference GMM and two-step system GMM techniques from dynamic panel data estimators. In addition to these factors, the effects of new factors on renewable energy were also investigated. Yang and Park (2020) discussed the effect of renewable energy financial incentive policy and democratic governance on increasing renewable energy efficiency by the panel data of 98 developed countries between 2000-2014 and the fixed-effects regression estimates (FE) method. Sasmaz et al. (2020) concluded that there was a directly proportional relationship between renewable energy consumption and human development index (HDI). With a different approach, Agyeman et al. (2020) investigated the long-term driving forces of renewable energy consumption by using a time series data from 1990 to 2015 and the Fully modified ordinary least square (FMOLS) estimator. Przychodzen and Przychodzen (2020) discussed the effect of widely used factors, such as GDP, FDI and CO_2 on renewable energy consumption. In addition, the factors of unemployment level, inflation, domestic credit availability, public debt, current account balance, R&D expenditures, corporate restructuring, competition policy implementation, energy market competitiveness, European Union membership, and being a party to the Kyoto Protocol were also examined.

It is seen in the literature that the studies were mostly carried out on countries. There are few studies discussing the regional differences of a country. Dorrell and Lee (2020) discussed the effect of political parties on wind energy in the United States of America regarding the states. Hartley et al. (2015) revealed the impact of alternative energy sources on local employment in the state of Texas by First difference GMM methods. Bao and Xu (2019) examined the geographical regions in China by utilizing the annual data of 30 provinces (excluding Tibet, Hong Kong, Macau and Taiwan) between 1997-2015 and concluded that the value of China's renewable energy consumption, urbanization, and economic growth in one region can be influenced by others.

GDP and population, used in the study as the determiners of renewable energy consumption, have also been subjected in most of the studies detailed above. However, the literature examining direct public investments (PI) as a factor effective on consumption is not comprehensive. Financial investments made by public and private sectors are of paramount importance to increase renewable energy generation capacity

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in leading economies (Lee, 2019). Policy-makers may provide funds to local authorities in order to be used in renewable energy development (Menz and Vachon, 2005). However, public funds are limited, and many renewable energy projects compete for these limited funds (Wang Q. et al., 2019). Polzin et al. (2015) investigated the effect of electricity generation capacity investments, made by institutional investors in accordance with public policy in OECD countries, on renewable energy (RE), and stated that there was generally a directly proportional effect. By conducting an analysis with the 2004-2015 panel data of 23 OECD countries, Corrocher and Cappa (2020) found a directly proportional relationship among public investments and public policies with solar energy.

The above empirical studies show to what extent there are few regional study results on the determinants of the consumption amount of renewable energy sources. In Turkey, regional development differences are high. Renewable energy sources are considered to be an opportunity to reduce these differences. In this respect, it is of importance that energy policymakers and governments should be aware of renewable energies. Therefore, the motivation for this study is to investigate the effect of regional GDP, public investments and population on renewable energy sources in Turkey. The results of this study will provide policy-makers with suggestions in creating sustainable energy policies for Turkey.

MAIN FOCUS OF THE CHAPTER

In this study, in order to reveal the regional differences in Turkey, 26 regions were evaluated in the group of Degree 2 according to Nomenclature of Territorial Units for Statistics (NUTS) (Table 1). This study is also important because of its contribution to regional data collection. Turkey was compiled by the author without using any database ready for data to be revealed regional differences. Regional renewable energy consumption and public investments data were compiled and shared for use in other research.

Dataset and Method

Regional data were used to investigate the regional impact on the consumption of renewable energy sources. All these are annual data related to 2010-2018 for 26 regions. The explanatory variables creating the determinants of the renewable energy consumption amount were specified by examining both accessibility and studies in the literature. In the study, the natural logarithms of all variables were taken and used. The sources of the explained and explanatory variables and data sets used in the study are provided in Table 2.

The reason for using the above-listed variables in our research model may be summarized as follows:

Regional GDP: This refers to the natural logarithmic form of the per capita GDP of the regions. As income increases, energy consumption increases, and GDP is accepted to be a factor affecting renewable energy consumption (Padhan et al., 2020). The regions with high value in this variable may better subsidize construction and preparation costs (Dorrell and Lee,2020).

Regional Population: This represents the natural logarithmic form of the total population of each region. As energy demand will be higher in the regions with higher populations, such regions need more energy generation to meet this demand. These regions are more encouraged to make an investment on alternative energy sources (Dorrell and Lee,2020).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Statistical Regional Units Classification (NUTS-2nd Level)

Region Code	Citys
TR10	İstanbul
TR21	Tekirdağ, Edirne, Kırklareli
TR22	Balıkesir, Çanakkale
TR31	İzmir
TR32	Aydın, Denizli, Muğla
TR33	Manisa, Afyon, Kütahya, Uşak
TR41	Bursa, Eskişehir, Bilecik
TR42	Kocaeli, Sakarya, Düzce, Bolu, Yalova
TR51	Ankara
TR52	Konya, Karaman
TR61	Antalya, Isparta, Burdur
TR62	Adana, Mersin
TR63	Hatay, Kahramanmaraş, Osmaniye
TR71	Kırıkkale, Aksaray, Niğde, Nevşehir, Kırşehir
TR72	Kayseri, Sivas, Yozgat
TR81	Zonguldak, Karabük, Bartın
TR82	Kastamonu, Çankırı, Sinop
TR83	Samsun, Tokat, Çorum, Amasya
TR90	Trabzon, Ordu, Giresun, Rize, Artvin, Gümüşhane
TRA1	Erzurum, Erzincan, Bayburt
TRA2	Ağrı, Kars, Iğdır, Ardahan
TRB1	Malatya, Elazığ, Bingöl, Tunceli
TRB2	Van, Muş, Bitlis, Hakkari
TRC1	Gaziantep, Adıyaman, Kilis
TRC2	Şanlıurfa, Diyarbakır
TRC3	Mardin, Batman, Şırnak, Siirt

Table 2. Variables Used in the Study

Variables	Description	Shortening	Source
Dependent variable	Energy generation from renewable sources (kWh)	lny	EPDK
Independent variable	Regional gross domestic product, by branches of economic activity (Thousand TL, at Current Prices)	lngdp	TUIK
Independent variable	Population(Thousand)	lnn	TUIK
Independent variable	Public investment program initial grant (Thousand TL, at Current Prices)	lnky	SBB

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Regional Public Investments: This refers to the natural logarithmic form of regional public investments. Energy investments from public investments are one of the physical infrastructure investment items such as roads, water and dams (Pirili, 2011). Domestic investments generally promote development and contribute to economic growth. In this respect, renewable energy sector can increase the preference of renewable energy use (Erdinş and Aydınbaş, 2020). In Turkey, there are large differences between regions in terms of development. The public investments are an important factor in overcoming these regional differences (Öztürk, 2012). The public investments made on a regional basis are also transferred to energy investments by development agencies. In this regard, the regional public investments were also included in this model. The descriptive statistics of the variables are given in Table 3.

Table 3. Descriptive Statistics

	LNY	LNGDP	LNKY	LNN
Mean	16.37	17.73	13.74	14.71
Median	19.23	17.68	13.68	14.76
Max	22.80	20.87	16.30	16.53
Min	0.00	15.90	12.57	13.51
Std. Dev.	7.40	0.91	0.66	0.59
Skewness	-1.65	0.71	1.05	0.53
Kurtosis	4.00	3.90	4.78	4.46

In this study, dynamic panel data analysis methods were used to explain the determinants of the amount of regional energy consumption based on renewable energy. GMM estimators provide effective results in cases where the time interval is small but cross-section sizes are large (Roodman, 2009).

Dynamic Panel Data Analysis

In the model established in this study, the effect of regional population, gdp and public investments on regional renewable energy consumption amount is investigated. As it is highly possible that there is a non-linear relationship between the variables used in our model, natural logarithmic forms of the variables were used when establishing the model.

Dynamic panel data analyses are widely used among panel data analysis methods. The dynamic relationships related to economic and financial variables are characterized by the existence of lagged values of the dependent variable along with independent variables (Demirci, 2018). In dynamic panel data, the cause and effect relationship of the underlying case is generally dynamic in time. For example, performance may be affected by previous year's expenses, which can play a more important role, rather than the current year's marketing expenses. In order to control this internal relationship, dynamic panel data estimation techniques also use the lags of dependent variables as explanatory variables (Ullah et. al, 2018). The estimators used in dynamic panel data analysis are PooledOLS (POLS), FixedEffect (FE), one-step difference GMM, two-step difference GMM, one-step system GMM and two-step system GMM.

While difference GMM estimator is suggested by Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998) recommend system GMM estimators. Blundell and Bond (1998) ar-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

gue that System GMM estimator provides consistent and efficient parameter estimations, and has better asymptotic and finite sampling properties than the simple first differences GMM estimator.

Estimations were performed by considering the dynamic panel features with lagged value of renewable energy consumption in order to determine the proper estimator for our model while estimator of difference and system generalized method of moments (GMM) was utilized in order to estimate our empirical model. Dahlberg and Johansson (2000) state that GMM techniques are the most proper methods that may be used when examining regional that is local administration behaviors. The GMM estimator produces efficient estimations in which the first differences of instrument variables are independent of fixed effects (Roodman, 2009). The GMM estimator generates consistent parameter estimates in a regression, where independent variables are not absolutely exogenous, and also controls heteroscedasticity and autocorrelation (Dorrell and Lee, 2020). In order to test the validity of instruments in our GMM estimation, serial correlation tests suggested by Arellano and Bover (1995) were administered (the null hypothesis: there is no autocorrelation), and it was tested for a series of quadratic correlations in residues. In addition, the restrictions in the estimators were examined by over identification Hansen test (null hypothesis: the overall validity of the instruments).

Model

The model that will reveal the relationship among the variables in this study is provided in Equation 1.

$$lnY_{it} = \beta_0 + \beta_1 lnY_{it-1} + \beta_2 lnGDP_{it} + \beta_3 lnKY_{it} + \beta_4 lnN_{it} + \epsilon_{it} \quad (1)$$

Y_{it} : Regional renewable energy consumption, Y_{it-1} : A lagged value of the regional renewable energy consumption amount, GDP_{it} : Regional economic growth, N_{it} : Regional population, KY_{it} : Regional public investments. What is more, ϵ_{it} refers to error term in the model. STATA package program was used to estimate this model.

SOLUTIONS AND RECOMMENDATIONS

The estimation results obtained by dynamic panel data analysis methods are given in Table 4.

When examining the estimation results in Table 4, it is concluded that the instruments used in the model are not over identified as null hypothesis was accepted in Hansen test statistics for both difference GMM and System GMM estimators. By this result, it is concluded that the estimation methods used are proper. When autocorrelation tests of difference GMM and System GMM estimation results are examined, there is a first order negative autocorrelation (AR1) in error terms while there is not any second order autocorrelation (AR2). Moreover, Sargan test p values in the GMM system are above the significance level, and confirm the validity of the study models and the instruments used. The study concluded that both differents GMM and system GMM techniques generated effective results. When the first lag coefficients of the dependent variable are compared, it is lower in the FE estimator (0,128) than in the POLS estimator (0,389). According to Bond's (2002) suggestion on the selection of the GMM type, as the ranking between the initial lag coefficients was concluded to be FE <GMM <POLS, difference-GMM is enough, thus, the use of System GMM is not very necessary.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Dynamic Panel Data Analysis of Regions

Variables	POOLED OLS	FIXED	one-step difference GMM	two-step difference GMM	one-step system GMM	two-step system GMM
lny L1	0.389*	0.128**	0.242*	0.285*	0.269**	0.280*
	(0.000)	(0.045)	(0.000)	(0.000)	(0.014)	(0.010)
lngdp	1.875***	8.543*	6.271***	6.923*	5.5189*	5.159*
	(0.054)	(0.001)	(0.103)	(0.000)	(0.000)	(0.002)
lnky	1.126	2.796**	2.393	2.002*	1.778***	1.727**
	(0.230)	(0.025)	(0.139)	(0.000)	(0.051)	(0.044)
lnn	-1.896	-25.527	-13.922	-32.394*	-7.011*	-6.631*
	(0.154)	(0.253)	(0.735)	(0.007)	(0.000)	(0.001)
_cons	-9.630	200.595			-6.277	-4.978
	(0.330)	(0.495)			(0.614)	(0.726)
Sargan Test	0.364	0.3501	32.33*	32.33*	38.77*	38.77*
Hansen J-test				22.310	24.760	24.760
				(0.173)	(0.211)	(0.211)
AR(1)			-6.32*	-2.43**	-2.75*	-2.14**
			(0.000)	(0.015)	(0.006)	(0.033)
AR(2)			1.01	0.78	0.84	0.76
			(0.313)	(0.434)	(0.403)	(0.446)

*, **, *** significant at 1%, 5%, 10% significance level, respectively. Over-identification for restrictions in GMM estimation is investigated by Hansen J-test. The first differences of residues are investigated by AR2 test of Arellano-Bond for the presence of second order autocorrelation.

When the results of the estimation are examined, the increase in per capita national income (with an average of 5.5% effect for a 1% increase in GDP) affects renewable energy consumption significantly and proportionally. While the regions with low income levels carry out infrastructure investments, the regions with high income make more investment in the development of renewable energy sources. In 2018, while the energy generation based on renewable sources was 1463.7 GWh in the region of TR10, having the highest economic growth in Turkey, this was 374.5 GWh in the region of TRA2 with the lowest figures. This result is in accordance with the conclusions achieved by Khan Z. et al. (2020), Omri et. al (2015), Agyeman et al. (2020), Przychodzen and Przychodzen (2020).

The population variable has a negative coefficient in all estimation results, and the coefficient of population is significant in System GMM models. This result is in the parallel of the direction of relationship specified by Dorrell and Lee (2020) in the study on the relationship between installed power of wind energy sources and population. Mehrara et al. (2015) stated that urban population is one of the most important variables that adversely affect renewable energy consumption in the Economic Cooperation Organization (ECO) economies. Agyeman et al. (2020) also suggested that the population has an adverse but significant effect on renewable energy consumption. This may refer to the contribution of working group population to economic growth and the increasing need for fossil sources, the primary source of energy generation (Zhao Y. et. al., 2013).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The directly proportional effect of public investments on renewable energy consumption shows that increasing public investments will increase the consumption of renewable energy. Transfer of a certain part of public investments to this sector will increase the spreading speed of private entrepreneurs regarding renewable energy investments, and will establish a safe environment for investment. The direction regarding the effect of public investments achieved in our results is in accordance with the result suggested by Le (2019).

FUTURE RESEARCH DIRECTIONS

The most significant contribution of this study is that it provides the first assessment results obtained by regional data regarding Turkey. In further studies, in order to discuss these factors in detail, the effects of them on hydroelectric, solar, wind, geothermal and biomass installed powers from different renewable sources, about which there is a high potential in Turkey, can be examined. The effects of these factors may be subjected for cities provided that required data are provided. In addition, the effect of different incentive implementation, in which YEKTEM regulation, used to promote renewable energy consumption in Turkey, is also included, may also be discussed.

CONCLUSION

The influence levels of different factors affecting renewable energy consumption can be analyzed with different studies. This study is limited to three factors: regional economic growth, public investments and population. This is because these factors are related to regional development from different angles. Regional economic growth is included in the study in terms of encouraging renewable energy consumption, regional public investments, since most of the investments in renewable energy generation facilities require government support, and lastly because of the potential of creating employment for the population, they are included in the study in terms of monitoring the migration situation.

Dynamic panel data analyses are frequently used for comprehensive analysis of these effects. This study concluded that both differents GMM and system GMM techniques have provided effective results. In all dynamic panel data analysis methods, the directions of the relationships between variables have been same. In this study, it has been found out that one-step system GMM and two-step system GMM techniques from dynamic panel data analysis methods have had significant values in all variable coefficients. It has been suggested according to the estimation results that regional economic growth and public investments affect the consumption of regional renewable energy resources in the same direction while regional population has an adverse effect.

There are some limitations in this study. The first one is the difficulty of compiling regional data. While the data for more than 20 years have been obtained for some of the variables, the period has been limited for some variables. For this reason, this study is limited to the period between 2010-2018. Another limitation is the exclusion of hydraulic power plants with dams, which are not included in the YEKDEM lists but provide a large part of production. In Turkey, renewable energy investments have increased in recent years. However, this increase has not been seen at the same level in each region. Therefore, there has been no production in some of the sources in certain periods in the 26 regions that have been examined.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this study, evaluation foreign direct investments, which were not included due to the lack of regional data but used in the general literature, in addition to public investments, may result in efficient results. This evaluation may be extended with the energy sources obtained from the hydraulic power plants with dam. In the following years, it is considered that renewable energy production sector in Turkey will develop more, thus, an extension in the period of this study may affect the results of the study.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Agyeman, J. K., Ameyaw, B., Li, Y., Appiah-Kubi, J., Annan, A., Oppong, A., & Twumasi, M. A. (2020). Modeling the long-run drivers of total renewable energy consumption: Evidence from top five heavily polluted countries. *Journal of Cleaner Production*, *277*, 123292. doi:10.1016/j.jclepro.2020.123292

Amoah, A., Kwablah, E., Korle, K., & Offei, D. (2020). Renewable energy consumption in Africa: The role of economic well-being and economic freedom. *Sustainability Science*, *10*(32), 2–17.

Anton, S. G., & Nucu, A. E. A. (2020). The effect of financial development on renewable energy consumption. A panel data approach. *Renewable Energy*, *147*, 330–338. doi:10.1016/j.renene.2019.09.005

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, *58*(2), 277–297. doi:10.2307/2297968

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51. doi:10.1016/0304-4076(94)01642-D

Assi, F., Isiksal, A. Z., & Tursoy, T. (2021). Renewable Energy Consumption, Financial Development, Environmental pollution, and Innovations in the ASEAN +3 group: Evidence from (P-ARDL) model. *Renewable Energy*, *165*(1), 689–700. doi:10.1016/j.renene.2020.11.052

Bao, C., & Xu, M. (2019). Cause and effect of renewable energy consumption on urbanization and economic growth in China's provinces and regions. *Journal of Cleaner Production*, *231*, 483–493. doi:10.1016/j.jclepro.2019.05.191

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143. doi:10.1016/S0304-4076(98)00009-8

Bond, S. R. (2002). *Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice.* Center for Microdata Methods and Practice Working Papers.

Byrne, J., Hughes, K., Rickerson, W., & Kurdgelashvili, L. (2007). American policy conflict in the greenhouse: Divergent trends in federal, regional, state, and local green energy and climate change policy, *Energy. Pol*, *35*(9), 4555–4573. doi:10.1016/j.enpol.2007.02.028

Corrocher, N., & Cappa, E. (2020). The Role of public interventions in inducing private climate finance: An empirical analysis of the solar energy sector. *Energy Policy*, *147*, 111787. doi:10.1016/j.enpol.2020.111787

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dahlberg, M., & Johansson, E. (2000). An examination of the dynamic behaviour of local governments using GMM bootstrapping methods. *Journal of Applied Econometrics*, *15*(4), 401–416. doi:10.1002/1099-1255(200007/08)15:4<401::AID-JAE564>3.0.CO;2-G

Demirci, N. S. (2018). The Firm Specific Determinants of Fixed Capital Investments: Theories of Corporate Investment and an Applıcatıon to Industrıal Firms Listed ın BIST. *Finance & Econometrics*, 111.

Dorrell, J., & Lee, K. (2020). The Politics of Wind: A state level analysis of political party impact on wind energy development in the United States. *Energy Research & Social Science*, *69*, 101602. doi:10.1016/j.erss.2020.101602

Energy Market Regulatory Authority (EPDK). (2021). *Nihai Yek Listeleri*. Retrieved from http: https://www.epdk.gov.tr/Detay/Icerik/372/elektrikyekdemhttp://www.epdk.org.tr/TR/Dokuman/8692. Erişim Tarihi 07/01/2021.

Erdinç, Z., & Aydınbaş, G. (2020). Panel Data Analysis on Determinants Of Renewable Energy Consumption, Journal Of Social. *Humanities and Administrative Sciences*, *6*(24), 346–358.

Güney, T. (2020). Renewable energy consumption and sustainable development in high-income countries. *International Journal of Sustainable Development and World Ecology*. Advance online publication. doi:10.1080/13504509.2020.1753124

Hartley, P., Medlock, K. B. III, Temzelides, T., & Zhang, X. (2015). Local employment impact from competing energy sources: Shale gas versus wind generation in Texas. *Energy Econ*, *49*, 610–619. doi:10.1016/j.eneco.2015.02.023

Instıtute, T. S. (TUIK). (2021). Retrieved from http: https://data.tuik.gov.tr/

International Council for Local Environmental Initiatives (ICLEI). (2021). *Climate Protection*. Retrieved from http: https://iclei-europe.org/our-members/

International Energy Agency (IEA). (2021). *Data and statistics.* Retrieved from http: https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20supply&indicator=TPESbySource

Khan, H., Khan, I., Oanh, L. T. K., & Lin, Z. (2020). The Dynamic Interrelationship of Environmental Factors and Foreign Direct Investment: Dynamic Panel Data Analysis and New Evidence from the Globe. *Mathematical Problems in Engineering*, 1–12.

Khan, Z., Malik, M. Y., Latif, K., & Jiao, Z. (2020). Heterogeneous effect of eco-innovation and human capital on renewable & non-renewable energy consumption: Disaggregate analysis for G-7 countries. *Energy*, *209*, 118405. doi:10.1016/j.energy.2020.118405

Lea, T.-H., Nguyen, C. P., & Park, D. (2020). Financing renewable energy development: Insights from 55 countries. *Energy Research & Social Science*, *68*, 101537. doi:10.1016/j.erss.2020.101537

Lee, T. (2019). Financial investment for the development of renewable energy capacity. *Energy & Environment*, *0*(0), 1–14. doi:10.1177/0958305X19882403

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mehrara, M., Rezaei, S., & Razi, H. D. (2015). Determinants of renewable energy consumption among ECO countries; based on Bayesian model averaging and weighted-average least square. *Int Lett Soc Hum Sci*, *54*, 96–109. doi:10.18052/www.scipress.com/ILSHS.54.96

Menz, F. C., & Vachon, S. (2006). The effectiveness of different policy regimes for promoting wind power: Experiences from the states. *Energy Policy*, *34*(14), 1786–1796. doi:10.1016/j.enpol.2004.12.018

Omri, A., Daly, S., & Nguyen, D. K. (2015). A robust analysis of the relationship between renewable energy consumption and its main drivers. *Applied Economics*, *47*(28), 2913–2923. doi:10.1080/00036846.2015.1011312

Öztürk, L. (2012). Public Investments and Regional Inequality: A Causality Analysis, 1975-2001. *Ege Academıc Revıew*, *12*(4), 487–495.

Padhan, H., Padhang, P. C., Tiwari, A. V., Ahmed, R., & Hammoudeh, S. (2020). Renewable energy consumption and robust globalization(s) in OECD countries: Do oil, carbon emissions and economic activity matter? *Energy Strategy Reviews*, *32*, 100535. doi:10.1016/j.esr.2020.100535

Pata, U. K. (2018). The influence of coal and noncarbohydrate energy consumption on CO2 emissions: Revisiting the environmental Kuznets curve hypothesis for Turkey. *Energy*, *160*, 1115–1123. doi:10.1016/j.energy.2018.07.095

Pirili, M. U. (2011). The Role of Public Investments in Regional Development: A Theoretical Review. *Ege Academıc Revıew*, *11*(2), 309–324.

Polzin, F., Migendt, M., Taube, ¨. F. A., & von Flotow, P. (2015). Public policy influence on renewable energy investments—A panel data study across OECD countries. *Energy Policy*, *80*, 98–111. doi:10.1016/j.enpol.2015.01.026

Przychodzen, W., & Przychodzen, J. (2020). Determinants of renewable energy production in transitioneconomies: A panel data approach. *Energy*, *191*, 116583. doi:10.1016/j.energy.2019.116583

Resıdency of the Republıc of Turkey Strategy and Budget Dırectorate (SBB). (n.d.). Retrieved from http: http://www.sbb.gov.tr/wp-content/uploads/2019/04/Yat%C4%B1r%C4%B1mlar%C4%B1n-%C4%B0llere-G%C3%B6re-Da%C4%9F%C4%B1l%C4%B1m%C4%B1.xlsx

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, *9*(1), 86–136. doi:10.1177/1536867X0900900106

Sasmaz, M. U., Sakar, E., Yayla, Y. E., & Akkucuk, U. (2020). The Relationship between Renewable Energy and Human Development in OECD Countries: A Panel Data Analysis. *Sustainability*, *12*(18), 7450. doi:10.3390u12187450

Turkish Renewable Energy Resources Support Mechanism (YEKDEM). (2011). *Official Gazette of the Republic of Turkey, 28782.* Retrieved from http: https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=18907&MevzuatTur=7&MevzuatTertip=5

Ullah, S., Akhtar, P., & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*, *71*, 69–78. doi:10.1016/j.indmarman.2017.11.010

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wang, B., Yang, Z., Xuan, J., & Jiao, K. (2020). Crises and opportunities in terms of energy and AI technologies during the COVID-19 pandemic. *Energy AI*, *1*, 100013. doi:10.1016/j.egyai.2020.100013

Wang, Q., Li, S., & Pisarenko, Z. (2020). Heterogeneous effects of energy efficiency, oil price, environmental pressure, R&D investment, and policy on renewable energy – evidence from the G20 countries. *Energy*, *209*, 118322. doi:10.1016/j.energy.2020.118322

Wang, Q., Valchuis, L., Thompson, E., Conner, D., & Parsons, R. (2019). Consumer Support and Willingness to Pay for Electricity from Solar, Wind, and Cow Manure in the United States: Evidence from a Survey in Vermont. *Energies*, *12*(23), 4467. doi:10.3390/en12234467

Yang, S., & Park, S. (2020). The effects of renewable energy financial incentive policy and democratic governance on renewable energy aid effectiveness. *Energy Policy*, *145*, 111682. doi:10.1016/j.enpol.2020.111682

Zhao, P., Lu, Z., Fang, J., Paramati, S. R., & Jiang, K. (2020). Determinants of renewable and non-renewable energy demand in China. *Structural Change and Economic Dynamics*, *54*, 202–209. doi:10.1016/j.strueco.2020.05.002

Zhao, Y., Tang, K.-K., & Wang, L.-L. (2013). Do renewable electricity policies promote renewable electricity generation? Evidence from panel data. *Energy Policy*, *62*, 887–897. doi:10.1016/j.enpol.2013.07.072

ADDITIONAL READING

Abdel-Basset, M., Gamal, A., Chakrabortty, R. K., & Ryan, M. J. (2021). Evaluation approach for sustainable renewable energy systems under uncertain environment: A case study. *RENEWABLE ENERGY, Volume*, *168*, 1073–1095. doi:10.1016/j.renene.2020.12.124

Alkan, O., & Albayrak, O. K. (2020). *Ranking of renewable energy sources for regions in Turkey by fuzzy entropy based fuzzy COPRAS and fuzzy multimoora, renewable energy*. Volume.

KEY TERMS AND DEFINITIONS

Economic Growth: Economic growth is Gross National Product (GNP) or Gross Domestic Product (GDP).

Energy Production: We need energy at every stage of our lives. For a country, the level of economic development is one of its indicators, and the amount of production due to energy consumption.

Population: All those who live or live in a country, a region, a city or a village.

Public Investment: Investments made by the state to increase the capital stock of the economy.

Regional Development Degrees: Regions with different shares of national income.

Renewable Energy: They are sustainable energy resources obtained from natural resources.

Renewable Energy Production: Energy obtained from renewable sources such as sun, wind, biomass.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 3
Reflection

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 24
Education for the Digital Industry:
Opportunities and Challenges of Experience-Based Expertise and Open Innovation

Valentina Chkoniya
https://orcid.org/0000-0003-1174-3216
University of Aveiro, Portugal

Fernando Cruz Gonçalves
ENIDH, Portugal

Maria Manuela Martins Batista
ENIDH, Portugal

ABSTRACT

Education holds the power to transform and enrich the lives of people. In the era of the digital industry, where data is omnipresent in every walk of life and new trends impact society and future jobs, humans continue to evolve through education and developing mechanisms to improve education with data science in the heart of it. This chapter demonstrates that experience-based expertise and open innovation must be understood as a single process, where living labs that involve academies and enterprises create unique conditions for society's progress. There is a trinomial relationship between academy, society, and industries, which are interestingly far more exploited than the education and research. Effective management of the knowledge and information transferred between open innovation ecosystem partners is crucial. The scientific development of both concepts is an active field in the academic community, and new ideas appear, opening new paths of knowledge transfer methods with knowledge from data.

DOI: 10.4018/978-1-7998-6985-6.ch024

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Education is the key to shaping the lives of people. It holds power to transform and enrich it. Since the dawn of civilization, humans have evolved through education and have developed mechanisms to improve education. In the 21st century, where data is omnipresent in every walk of life, education is no exception.

Due to the recent explosion of big data, our society has been rapidly going through digital transformation and entering a new world with numerous eye-opening developments. These new trends impact society and future jobs, and thus student careers. At the heart of this digital transformation is data science (Braschler, Stadelmann, & Stockinger, 2019; Song & Zhu, 2017).

The three main differences between the fourth and previous three industrial revolutions are speed, scope, and systemic impact. Compared to earlier industrial revolutions, the fourth is developing exponentially rather than linearly (June Kim, 2021). Are the current education systems and paradigms oriented towards the values and goals of Industry 4.0? It does not always seem to be so. A 20th-century educational model is mostly still leading, introducing standardized facts and procedures designed to prepare the workforce for jobs that probably may not exist for a long time. It is not enough to cope with future challenges.

According to Vygotsky's Sociocultural Theory of Cognitive Development, "learning from experience is the process whereby human development occurs" (Vygotsky, Cole, John-Steiner, Scribner, & Souberman, 1978). Although experience is essential to learning, it is not enough; one has to do something to construct knowledge. Building an environment where students feel free to expose their knowledge state and compassionately help each other learn is incredibly difficult (Chkoniya, 2021). The idea of a growth mindset was first described by Carol Dweck (2017). Their research points out the importance of students believing that they can improve through practice. More interestingly, perhaps, they offer suggestions as to how to foster growth mindsets, such as praising effort instead of skill (Dweck, 2019; Hicks & Irizarry, 2018).

Business operations are undergoing drastic changes due to the disruptive effects of technological innovations, and it has consequences in the education sector (Oke & Fernandes, 2020). There are two essential dimensions of the learning process: grasping and transforming experience. A firm's capacity for accumulation is its most crucial technological asset. Therefore, the following virtuous cycle should be established: first, practical knowledge is accumulated systematically in the workplace; second, the knowledge should be scientifically codified. Third, the documentation of this codified knowledge should be conducted so those good textbooks can be produced. Fourth, training and teaching methods should be developed and tested in terms of their effectiveness and validity. Fifth, the new practices established on the foundation of the newly codified knowledge should be tried, and the performance of these trials should be evaluated objectively. Then, this cycle should be repeated, starting with the first phase (Lee et al., 2018).

This chapter demonstrates and gives an example that Experience-Based Expertise and Open Innovation (OI) must be understood as a single process. It is essential to create knowledge and transfer it, ensuring that the flow of knowledge takes place in a bidirectional way.

Background

Since Klaus Schwab and the World Economic Forum declared the arrival of the Fourth Industrial Revolution, there has been much discussion about it. The so-called fourth industrial or digital revolution is characterized by the unification of technologies that breaks the boundaries between physical, digital,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

and biological disciplines (Schwab, 2017). Compared to the three previous industrial revolutions, the main differences of fourth are speed, extent, and systemic impact (Chkoniya, Madsen, & Bukhrashvili, 2020). Only jobs that machines cannot do will last for the people in the future; they will base on creative expression, social interaction, physical dexterity, empathy, ingenuity, and collaboration (Lee et al., 2018). Lifelong inclusive and equitable education, formal and informal, physical and digital, will be vital for preparing new populations and society to be successful in this unpredictable future. (Alop, 2019).

Since the expected changes are disruptive, it will be difficult to adequately prepare for the changes to come with traditional high education (Alop, 2019; Erdogan & Demirel, 2017; Lee et al., 2018). The process of creating, approving, and launching new curricula, modules, subjects are now extended, cumbersome, and unable to respond quickly to increasingly rapid changes. Even the renewal of an existing curriculum may take a year or even more; creating new curricula maybe even longer and more cumbersome, and bureaucratic. These new curricula or their innovations may already be outdated when they start working.

Through the years, the education sector has been influenced by the evolution of technology and the global, regional, and local social dynamics. Therefore, today, teaching and learning methods include emerging information and communication technologies (ICTs), advanced tools, and innovative facilities. Also, teaching and learning methods are focused on developing desirable competencies in today's students to motivate them to propose solutions to today's and future problems. This new era is known as "Education 4.0". The OI Laboratories are used as enabling resources to reach the vision of Education 4.0 and boost knowledge transfer (Miranda et al., 2019).

MAIN FOCUS OF THE CHAPTER

It would be more effective to establish Enterprises and Academies, exchanging Experience-Based Expertise and OI. Especially when the massive digitalization of the Industry becomes evident, changing everything from the types of jobs available to the skills we will need.

Experience Based Expertise

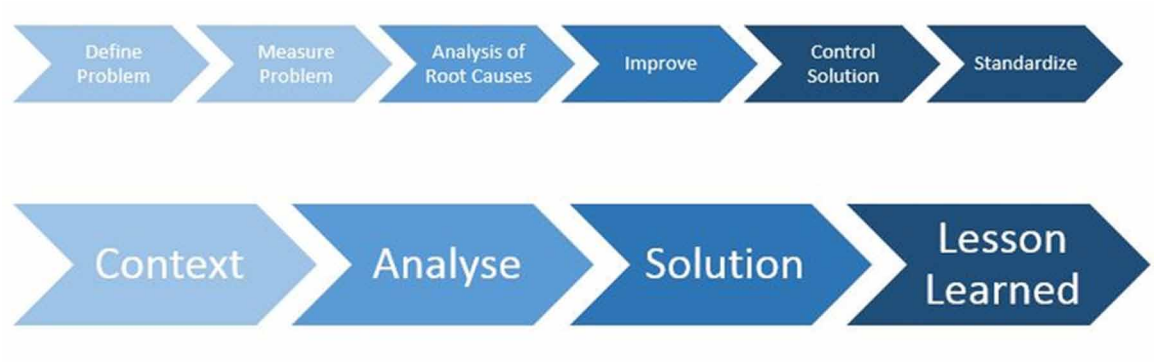
Learning is the process whereby knowledge is created through experience transformation (Kolb & Kolb, 2017). Experiential learning, which has its origin in the works of Dewey, Lewin, and Piaget, focuses on the central role that experience plays in the learning process (Holzer & Andruet, 1995). Experience-based knowledge plays an essential role in decision-making processes as decision-makers try to compare the current problems with previous situations to find an appropriate solution. Therefore, generating knowledge from past decisions can significantly facilitate and accelerate future decisions (Argote & Ingram, 2000; Knowledge Management, 2021). According to Matthew and Sternberg, experience-based knowledge is context-dependent and typically develops over time through an iterative learning process of perception, action, and feedback (Matthew & Sternberg, 2009).

The Industry has introduced various problem-solving methods that apply this concept of cognitive experience. One famous example is called Six-Sigma (Figure 1). It was created in the late 1970s, when cheap Japanese goods were becoming increasingly competitive on the global markets, forcing American companies to seek ways to improve their product quality while, at the same time, keeping the costs low. In a joint effort, General Motors, Ford, and Chrysler made large-scale use of SPC - Statistical Process Control, which enabled them to steadily track the manufacturing process at critical points of quality and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

costs. Similarly, Motorola, which was being driven out of the market by cheaper and superior Japanese products, gathered a group of distinguished experts in mathematical statistics, design, and quality assurance, who worked together to develop a consistent system for constant quality improvement. Today Six Sigma is an economic process allowing enterprises to improve their financial health through planning and controlling the workflow in such a way as to minimize the consumption of raw materials and the production of waste while at the same time providing more satisfaction for the customer (Kęsek, Bogacz, & Migza, 2019).

Figure 1. Six-Sigma as an implementation of the general experience feedback process (adapted from Jabrouni, Kamsu-Foguem, Geneste, & Vaysse, 2011)



As a common good, open science opens possibilities for the development of nations through innovations and collaborative constructions, which help to democratize knowledge (Ramírez & García-Peñalvo, 2018).

Knowledge transfer in the context of Academia-Enterprise collaboration is understood as a determinant that improves innovation and competitiveness. The mechanisms by which this collaboration occurs have awakened increasing interest among academics, as well as in the business sector (Vélez-Rolón, Méndez-Pinzón, & Acevedo, 2020). The most recognized source of knowledge used in new research is directed towards Academia-Enterprise relationships in an OI environment (Alvarez-Meaza, Pikatza-Gorrotxategi, & Rio-Belver, 2020).

Open Innovation

Innovation is a crucial aspect of the sustainability and competitiveness of organizations and the economy. In recent years, OI has burst onto the scene as a more open approach, in which the sources and exchange of knowledge are extended beyond organizational borders. One of the main actors and providers of knowledge is the Academy (Heras-Rosas & Herrera, 2021).

The Fourth Industrial Revolution can be recognized as the advancement from simple digitization to innovation based on combinations of technologies enabling companies to innovate their business models (Lee et al., 2018). Companies use innovation to generate and apply knowledge, develop new products and services, new business models, and ultimately obtain business results.

Traditionally, companies innovated only using internal resources. The main reasons were the non-commercial focus of the scientific community and their need to protect their valuable assets. Such

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

companies prioritized the volume of resources, the number of projects, and the investments in innovation. Consequently, organizations with closed innovation models tend to have big research departments generating plenty of in-house knowledge. The more, the better.

Today, the Digital Industry has undoubtedly become a key enabler of innovation, when innovation processes have become more open and require more significant resources in the different implementation phases to capture and transfer knowledge within and outside the Enterprise's boundaries (Urbinati, Chiaroni, Chiesa, & Frattini, 2020). OI ecosystems involve the transfer of knowledge between multiple stakeholders to contribute toward product and service innovation, and to an extent, have superseded network-level approaches to co-creation. Effective management of the knowledge and information transferred between ecosystem partners is crucial for the process of OI (Bacon, Williams, & Davies, 2019). The OI has emerged recently as a new innovation model and already became a hot topic that currently facilitates establishing links between Academia and Industry (Neves, Costa, & Reis, 2021). It encourages companies to use the existing external knowledge rather than reinvent the wheel (Table 1).

Table 1. Comparison between open and closed innovation models (adapted from Ennomotive, 2020)

Characteristic	**Type of innovation model**	
	Closed Model	**Open Model**
Source of knowledge	Use knowledge from internal sources	Internal + External knowledge sources
Success rate	Low (typically 20-30%) success rate	Double innovation success rate, up to 80%
Productivity	A lot of rework, low productivity	40-60% productivity increase
Speed	Low speed of innovation	High speed of innovation (x 3)

It is favorable to establish links with other companies, academies, tech centers, and other knowledge sources. The term OI as we know it was coined by Henry Chesbrough, associate professor and head of the Open Innovation Center of the Haas Business School of the University of California. However, collaboration and co-creation have always been around, maybe since humans started trading or since the cottage industry exchanged practices from other regions and cultures (Bogers, Chesbrough, Heaton, & Teece, 2019; Dodgson, Gann, & Salter, 2006; Ennomotive, 2020; Howard, Steensma, Lyles, & Dhanaraj, 2016; Karhade & Dong, 2021):

1. Cisco may be the most remarkable OI pioneer. Its acquisition strategy allowed the company to grow during the late 90's. The goal of these acquisitions was to make hardware compatible with software in new technological products. This customer-driven innovation strategy helped Cisco over perform Bell Labs or Lucent by effectively translating significant tech investments into business growth.
2. Another trailblazer is, undoubtedly, the pharmaceutical company Eli-Lilly, spearheaded by its R&D strategy manager. The company reflected on using collective intelligence to improve the traditionally low innovation success rates in the pharmaceutical Industry. Consequently, e.Lilly was born around the year 2000, the first OI platform to connect with global scientific knowledge. This platform is the origin of Innocentive, an Eli Lilly spin-off whose goal was to allow other pharmaceutical and consumer goods companies to use this open model.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. The third initiator is Procter & Gamble. The company created in 2000 the so-called Connect & Develop their new OI process. P&G's ultimate goal was to continue growing by $2 billion annually while keeping a continuous R&D investment. Procter managed to have up to 50% of its innovations coming from the outside and, consequently, its R&D productivity increased by 60%.

OI has become a mainstream phenomenon in the current business landscape. Where two key innovation project attributes (i.e., complexity and uncertainty) are related to five factors for successful OI management (Bagherzadeh, Markovic, & Bogers, 2021):

1. Openness level;
2. External partner choice (e.g., users, suppliers, Academy, competitors, entrepreneurs, and start-ups);
3. OI mechanism choice (e.g., licensing agreements, alliances, innovation contests, and crowdsourcing intermediaries);
4. Collaboration process formalization;
5. Internal firm practices.

OI is improving innovation ecosystems in terms of promoting smart and responsible innovation. There is a trinomial relationship between Academy, Society, and Industries, which are interestingly far more exploited than the Education and Research. Given the noticeable importance of producing knowledge and disseminating that information, Academy will continue to have a prominent centrality in society in that regard (Neves et al., 2021).

Experience Based Expertise and OI must be understood as a single process. The scientific development of both concepts is an active field in the academic community, and also, that new terms appear, opening new paths of research (Alvarez-Meaza et al., 2020). This chapter gives a particular focus on Living Labs that involve Academies and Enterprises.

Living Labs

The "Living Lab" is an approach used since the 2000s in Europe and is defined as an OI laboratory that integrates the user as a central actor in the innovation process. By integrating users into the innovation process ensures highly reliable market assessment to reduce technological and commercial risks. "Living Lab" provides a better understanding of the user needs, which is a crucial factor for successful innovation and implementation of Pilot Projects (Arnould, Morel, & Fournier, 2020; Geenhuizen, 2018; Hossain, Leminen, & Westerlund, 2019; Leminen, Westerlund, & Nyström, 2012; Pallot & Pawar, 2012; Schuurman, De Marez, & Ballon, 2015).

Organizational usability must instead be evaluated in situ while the system is used for real work. There are three contexts for such evaluation: pilot implementation, techno change, and design in use (Table 3). The pilot implementation aims to inform the finalization of a system based on testing it in the field before go-live. Techno change focuses on shaking down a system during go-live to realize the benefits it was developed to help achieve. Design in use is the tailoring performed by users after go-live to fit a system and its use to their local and emerging needs (Hertzum, 2021).

According to Hertzum et al. (2021), a pilot implementation consists of five activities:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Characteristics of different types of living labs (adapted from Leminen, Westerlund, & Nyström, 2012)

Characteristic	Type of Living Labs			
	Utilizer-driven	**Enabler-driven**	**Provider-driven**	**User-driven**
Purpose	Strategic R&D activity with preset objectives	Strategy development through action	Operations development through increased knowledge	Problem-solving by collaborative accomplishments
Organization	Network forms around a utilizer, who organizes action for rapid knowledge results	Network forms around a region (regional development) or a funded project (e.g., public funding)	Network forms around a provider organization(s)	Network initiated by users lacks formal coordination mechanisms
Action	Utilizer guides information collected from the users and promotes knowledge creation that supports the achievement of preset goals	Information is collected and used together, and knowledge is co-created in the network.	Information collected for immediate or postponed use; new knowledge based on the information that the provider gets from the others	Information is not collected formally and builds upon users' interests; knowledge is utilized in the network to help the user community.
Outcomes	New knowledge for product and business development	Guided strategy change into a preferred direction	New knowledge supporting operations development	Solutions to users' everyday-life problems
Lifespan	Short	Short/medium/long	Short/medium/long	Long

1. Planning and design;
2. Technical configuration;
3. Organizational adaptation;
4. Pilot us;
5. Learning.

The first three activities are preparations. During the preparations, the focus of the pilot implementation is defined, the boundary between the pilot site and the organization at large is determined, the system is configured for the pilot site, interfaces to other systems are established, work procedures at the pilot site are adjusted, users receive training, safeguards against breakdowns are set up, and so forth. The extent and complexity of the preparations demonstrate that a pilot implementation is not just the period of pilot use. During the pilot use period, the pilot site staff has the double task of conducting their work with the system and providing input to the evaluation. The final activity, learning, is the reason for conducting pilot implementations (Figure 2). Importantly, learning about the organizational usability of the system occurs during the preparations and during the period of pilot use.

Example of the Education and Digital Shipping Industry

In the 21st century, the maritime Industry has become concerned about Maritime Education and Training, and it has been deemed as a positive and potentially game challenging strategy in order to solve

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Three Contexts for Evaluating Organizational Usability (adapted from Hertzum, 2018)

Time	Pilot implementation	Techno change	Design in use
	Before go-live	During go-live	After go-live
Definition	Pilot implementation is a field test of a properly engineered yet unfinished system in its intended environment, using actual data, and aiming – through real-use experience – to explore the value of the system, improve or assess its design, and reduce implementation risk.	Deliberate techno change is "the use of IT to drive improvements in organizational performance."	Design in use encompasses the "practices of interpretation, appropriation, assembly, tailoring and further development of computer support in what is normally regarded as deployment or use."
Purpose	To learn about the fit between the system and its context to explore the system's value, improve its design, and reduce implementation risk.	To transition from old practices to the new system and start realizing the benefits that motivated its introduction.	To appropriate—tailor—the system for local and emergent needs when opportunities for such appropriation are seen and seized
People	Usability practitioners plan and facilitate the process	Usability practitioners facilitate the process	Users drive the process, facilitated by usability practitioners
Setting	In the field, i.e., during real work, but limited to a pilot site	In the field, but while the system is still new and unfamiliar	In the field, during regular use of the system for real work
System	The pilot system, i.e., a properly engineered yet unfinished system	Finished system but not yet error-free and not yet fully configured	The finished system, yet with possibilities for reconfiguration
Process	Used in situ for a limited time and with special precautions against errors	Used in situ by inexperienced users, possibly with extra support during go-live	Used in situ by regular users, some of whom occasionally engage in design-in-use activities
Benefit focus	Specified benefits dominate; other benefits may emerge	Specified benefits dominate; other benefits may emerge	Emergent benefits are likely to dominate
Duration	Temporary, typically weeks or months	Temporary, typically months	Continuous, typically years or decades
Main challenge	The boundary between pilot site and organization at large	Premature congealment of the process to ensure benefit from the new system	Insufficient capability to make or disseminate changes

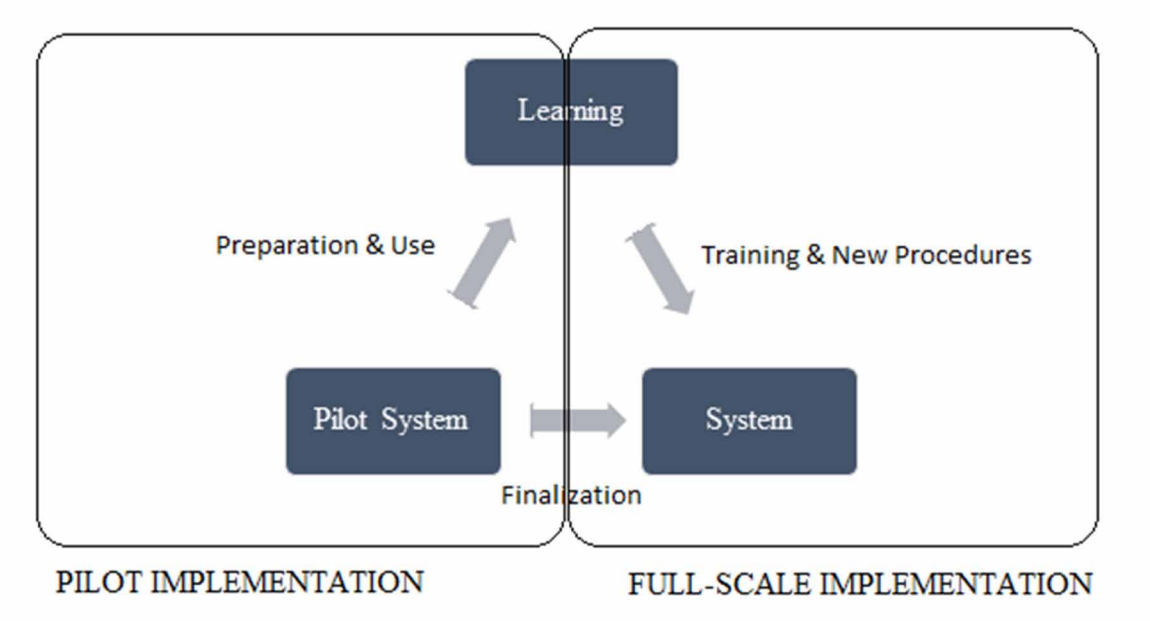
several actual and possible challenges that provide the enhancement of growth, expansion, and positive change (Basak, 2017).

As a result of improvements in information technology and infrastructure, maritime technologies and operations have dramatically evolved from traditional automatic, mechanical, mechatronic-based technologies to intelligence, human-centered, and in-formation and ICT-supported smart operations. Such changes subvert the traditional evaluation of the competence of individual laborers (Pan, Oksavik, & Hildre, 2020). At the same time, the difficulty of evaluating the organizational usability of those technologies is raised. The two are combined dramatically to challenge current maritime education and research on many levels (Hertzum, 2018; Pan et al., 2020).

It will be especially challenging in the future when the whole shipping process will become Smart. Smart shipping is a standard system where ships and ports are connected to one global system and operated in the most optimal way using standard algorithms. These systems require full or large-scale integration of Smart Ships and Smart Ports into a single network leading by so-called artificial intelligence. However, a human operator must be maintained. (Alop, 2019).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Pilot implementation (adapted from Hertzum, 2021)



To prepare students for jobs that have not been created yet, we need to teach them how to be great problem solvers to be ready for anything. One way to do this is by teaching content and skills using real-world case studies and a learning model focused on reflection during the problem-solving process (Heirs & Manuel, 2021; Schwartz, 2019). It is similar to the project-based learning and Growth Mindset idea, following the Theory of Cognitive Development (Alop, 2019; Dweck, 2019; Heirs & Manuel, 2021; Vygotsky et al., 1978).

Historical issues caused the gaps of maritime education in higher education and led to an unsystematic structure for research and development of maritime technology. Being an expert, or being at any stage of skill acquisition, does not necessarily mean performing and everyone else who exhibits the same type of thought processes (Basak, 2017). Everyone function in at least one of five stages of skill level:

1. Novice;
2. Advanced beginner;
3. Competent;
4. Proficient;
5. Expert.

In order to better prepare for the future, we need to include studying phenomena in a systematic way of participants in their everyday settings, taking a holistic view, providing a descriptive understanding, and taking a member's perspective. Therefore, there is no necessary to distinguish who is providing what types of maritime education, but we can see them as an utterly organizational system, including humans, technological artifacts, and institutional rules for organizing humans and technologies together (Pan et al., 2020).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Maritime education as a whole is generally considered repetitive practice-based training. There is a need for defining clear learning outcomes, improving the learning content to enable exploration and second-chance learning, minimizing theory-practice gaps by ensuring skills-knowledge balance and in-depth scholarship building, facilitating tasks for learning preparation and learning extension, and repositioning simulation components and their assessment schemes across the academic program (Jamil & Bhuiyan, 2021).

How to be more flexible and responsive to rapid changes? What are the new principles of education in the dynamically developing shipping industry and digital world where all information is available to everyone anytime? How to keep pace with the rapid and extensive changes in industries, including shipping? Lifelong inclusive and equitable education, formal and informal, physical and digital, will be vital for preparing new populations and society to be successful in this unpredictable future (Alop, 2019).

Hands-on experiential learning is a mainstay of meaningful education. Engaging students in field-based, experiential learning can help develop true passion and commitment to STEM (Science, Technology, Engineering, Math) subjects, allowing students to apply their skills and content knowledge in an authentic, experiential context (Living Labs) while enhancing many academic skills and engaging them in contextualized learning across disciplines (O'Neil et al., 2020).

We can find an interesting example in Portugal, where the only national college dedicated to the training of Merchant Marine Officers and senior managers of the Maritime-port sector in the areas of Intermodality, Logistics, and Management is participating in two pilot projects using an innovative approach in technological and pedagogical terms, through 'Living Labs' (ENIDH, 2021)

As a leading institution in the training of professionals linked to the maritime and logistics areas, the Escola Superior Náutica Infante D. Henrique (ENIDH) is involved in the eShip project, thus embodying its role as promoter of technological and digital progress in the sector. eShip project aims to digitize the vessel chartering process, thereby cutting management costs and fueling the digital transition.

The primary objective of the eShip project is to develop a digitalization model for the entire chartering process, from the market research stage to the completion of the trip, including freight negotiation and contracting, real-time trip analysis, overage calculations. /sub-stay, issuance of shipping and billing documents, post-trip analysis, and regulatory compliance. This electronic platform will be tested in two pilot projects using an innovative approach in technological and pedagogical terms, through Living Labs, focusing on the Liquefied Natural Gas (LNG) market, which is located in evident rise, and Solid Bulk.

SOLUTIONS AND RECOMMENDATIONS

Experience Based Expertise and OI must be understood as a single process, where Living Labs that involve Academies and Enterprises create unique conditions for society's progress. There is a trinomial relationship between Academy, Society, and Industries, which are interestingly far more exploited than the Education and Research. Effective management of the knowledge and information transferred between OI ecosystem partners are crucial.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

The world today is extremely well connected and heavily dependent on maritime trade. The efficiency of the maritime sector generally strengthens the country's international and political position, as its activities are linked to international economic cooperation, attracting foreign investment, membership in international organizations, and other vital factors (Agbaba, 2020). The shipping market is the result of global production and its constant changes. As part of such a market, maritime transport is vulnerable to external factors, such as international trade, political situations, financial trends, technological developments, and legislation, which may directly or indirectly affect demand in the sector or similar. The adverse effects of these changes are reflected in lower freight rates, lower daily rental prices, and reduced prices of new or used boats. The maritime crisis has significantly affected income, wages, number of employees, and similar. In order to ensure the development of the maritime sector, it is necessary to continue research with Living Labs that invest in the growth of transport capacities, deepen ports, build new terminals and modernize existing ones. It also requires open-innovation principles to co-create the advanced in transport development and application, technological, technical, economic, organizational, and commercial measures to adapt to the environment.

CONCLUSION

During the long history of higher education institutions, their value to society has suffered several changes. However, experience accumulated in enterprises is essential to learning. The essence of innovation lies in knowledge, so OI opens the door to knowledge transfer with agents outside the organization.

Experience Based Expertise and OI must be understood as a single process; that is to say, it is essential to create knowledge and be able to transfer it, ensuring that the flow of knowledge takes place in a bidirectional way. This process from the postulated model must end the generation of value for both organizations, innovation being a fundamental factor that must be promoted and measured in future investigations. Knowledge management and knowledge transfer in the Academia–Enterprise relationship can be interpreted as a single factor that mediates the flow and use of knowledge. (Vélez-Rolón et al., 2020). The three main differences between the fourth and previous three industrial revolutions are speed, scope, and systemic impact. Compared to earlier industrial revolutions, the fourth is developing exponentially rather than linearly (June Kim, 2021). OI has become a mainstream phenomenon in the current business landscape (Bagherzadeh et al., 2021), and the OI community is pivotal to the relationship between academia and Enterprise. To prepare students, we need to teach them how to be great problem solvers. One way to do this is by teaching content and skills using real-time connections with Enterprises. Since the expected changes are disruptive and understanding user needs a key factor for successful innovation, Living Labs create unique conditions for society's progress.

ACKNOWLEDGMENT

The eShip project is supported by the EEA Grants.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Agbaba, R. (2020). Maritime Challenges in Crisis Times. *Annals of Maritime Studies / Pomorski Zbornik, 59*(1), 51–60. . doi:10.18048/2020.59.03

Alop, A. (2019). The Challenges of the Digital Technology Era for Maritime Education and Training. *2019 European Navigation Conference (ENC), Navigation Conference (ENC), 2019 European*. 10.1109/EURONAV.2019.8714176

Alvarez-Meaza, I., Pikatza-Gorrotxategi, N., & Rio-Belver, R. M. (2020). Knowledge Sharing and Transfer in an Open Innovation Context: Mapping Scientific Evolution. *Journal of Open Innovation*, *68*(4), 1–23. doi:10.3390/joitmc6040186

Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, *82*(1), 150–169. doi:10.1006/obhd.2000.2893

Arnould, M., Morel, L., & Fournier, M. (2020). Developing a territorial diagnostic as part of a living lab process: Implementation to improve management and wood mobilization in small French private forest. *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Engineering, Technology and Innovation (ICE/ITMC), 2020 IEEE International Conference On*. 10.1109/ICE/ITMC49519.2020.9198373

Bacon, E., Williams, M. D., & Davies, G. H. (2019). Recipes for success: Conditions for knowledge transfer across open innovation ecosystems. *International Journal of Information Management*, *49*, 377–387. doi:10.1016/j.ijinfomgt.2019.07.012

Bagherzadeh, M., Markovic, S., & Bogers, M. (2021). Managing Open Innovation: A Project-Level Perspective. *IEEE Transactions on Engineering Management, Engineering Management, IEEE Transactions on*. *IEEE Transactions on Engineering Management*, *68*(1), 301–316. doi:10.1109/TEM.2019.2949714

Basak, S. K. (2017). A Framework on the Factors Affecting to Implement Maritime Education and Training System in Educational Institutions: A Review of the Literature. *Procedia Engineering*, *194*, 345–350. doi:10.1016/j.proeng.2017.08.155

Bogers, M., Chesbrough, H., Heaton, S., & Teece, D. J. (2019). Strategic Management of Open Innovation: A Dynamic Capabilities Perspective. *California Management Review*, *62*(1), 77–94. doi:10.1177/0008125619885150

Braschler, M., Stadelmann, T., & Stockinger, K. (2019). *Applied data science : lessons learned for the data-driven business. Applied data science lessons learned for the data-driven business*. Springer., doi:10.1007/978-3-030-11821-1

Chkoniya, V. (2021). *How Harvard's Case Method Can Help to Overcome the Challenges of Applied Data Science Education*. Academic Press.

Chkoniya, V., Madsen, A. O., & Bukhrashvili, P. (2020). *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior*. IGI Global. doi:10.4018/978-1-7998-3115-0

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

de las Heras-Rosas, C., & Herrera, J. (2021). Research Trends in Open Innovation and the Role of the University. *Journal of Open Innovation*, *7*(1), 1–22. doi:10.3390/joitmc7010029

Dodgson, M., Gann, D., & Salter, A. (2006). The role of technology in the shift towards open innovation: The case of Procter & Gamble. *R & D Management*, *36*(3), 333–346. doi:10.1111/j.1467-9310.2006.00429.x

Dweck, C. (2019). What Having a "Growth Mindset" Actually Means. *Harvard Business Review*.

Dweck, C. S. (2017). *Summary of Carol S. Dweck's Mindset: Key Takeaways & Analysis*. CreateSpace Independent Publishing Platform.

ENIDH. (2021). *Escola Superior Náutica Infante D. Henrique*. Retrieved from https://www.enautica.pt/en/

Ennomotive. (2020). *Open Innovation: Definition and Types of Innovation*. Author.

Erdogan, O., & Demirel, E. (2017). New Technologies in Maritime Education and Training, Turkish Experiment. *Universal Journal of Educational Research*, *5*(6), 947–952. doi:10.13189/ujer.2017.050606

Heirs, S., & Manuel, M. (2021). Sustainable Maritime Career Development: A case for Maritime Education and Training (MET) at the Secondary Level. *TransNav: International Journal on Marine Navigation & Safety of Sea Transportation*, *15*(1), 91–99. doi:10.12716/1001.15.01.08

Hertzum, M. (2018). Three contexts for evaluating organizational usability. *Journal of Usability Studies Archive*, *14*, 35–47.

Hertzum, M. (2021). *Organizational Implementation: The Design in Use of Information Systems*. Morgan & Claypool Publishers.

Hicks, S. C., & Irizarry, R. A. (2018). A Guide to Teaching Data Science. *The American Statistician*, *72*(4), 382–391. doi:10.1080/00031305.2017.1356747 PMID:31105314

Holzer, S., & Andruet, R. (1995). *A multimedia workshop learning environment for statics*. doi:10.1109/FIE.1995.483056

Hossain, M., Leminen, S., & Westerlund, M. (2019). A systematic review of living lab literature. *Journal of Cleaner Production*, *213*, 976–988. doi:10.1016/j.jclepro.2018.12.257

Howard, M., Steensma, H. K., Lyles, M., & Dhanaraj, C. (2016). Learning to collaborate through collaboration: How allying with expert firms influences collaborative innovation within novice firms. *Strategic Management Journal, 37*(10), 2092–2103. doi:10.1002mj.2424

Jabrouni, H., Kamsu-Foguem, B., Geneste, L., & Vaysse, C. (2011). Continuous improvement through knowledge-guided analysis in experience feedback. *Eng. Appl. of AI*, *24*(8), 1419–1431. doi:10.1016/j.engappai.2011.02.015

Jamil, M. G., & Bhuiyan, Z. (2021). Deep learning elements in maritime simulation programs: A pedagogical exploration of learner experiences. *International Journal of Educational Technology in Higher Education*, *18*(1), 1–22. doi:10.118641239-021-00255-0

June Kim, S. (2021). How can higher maritime education lead shipping growth? Korea's experience, 1948–1982. *International Journal of Maritime History*, *33*(1), 90–117. doi:10.1177/0843871420974062

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Karhade, P. P., & Dong, J. Q. (2021). Innovation Outcomes of Digitally Enabled Collaborative Problemistic Search Capability. *Management Information Systems Quarterly*, *45*(2), 693–717. doi:10.25300/MISQ/2021/12202

Kęsek, M., Bogacz, P., & Migza, M. (2019). The application of Lean Management and Six Sigma tools in global mining enterprises. *IOP Conference Series. Earth and Environmental Science*, *214*, 12090. doi:10.1088/1755-1315/214/1/012090

Knowledge Management. (2021). *Experience-Based Knowledge*. Knowledge management, Group Work, UCPori. Retrieved from https://knowledgemanagement5.wordpress.com/experience-based-knowledge/

Kolb, D., & Kolb, A. (2017). *The Experiential Educator: Principles and Practices of Experiential Learning*. Academic Press.

Lee, M., Yun, J. J., Pyka, A., Won, D., Kodama, F., Schiuma, G., Park, H. S., Jeon, J., Park, K. B., Jung, K. H., Yan, M.-R., Lee, S. Y., & Zhao, X. (2018). How to Respond to the Fourth Industrial Revolution or the Second Information Technology Revolution? Dynamic New Combinations between Technology, Market, and Society through Open Innovation. *Journal of Open Innovation*, *4*(3), 21. Advance online publication. doi:10.3390/joitmc4030021

Leminen, S., Westerlund, M., & Nyström, A.-G. (2012). Living Labs as Open-Innovation Networks. *Technology Innovation Management Review*, *2*(9), 6–11. doi:10.22215/timreview/602

Matthew, C. T., & Sternberg, R. J. (2009). *Developing experience-based (tacit) knowledge through reflection. In Learning and Individual Differences.* Elsevier Science., doi:10.1016/j.lindif.2009.07.001

Miranda, J., Lopez, C. S., Navarro, S., Bustamante, M. R., Molina, J. M., & Molina, A. (2019). Open Innovation Laboratories as Enabling Resources to Reach the Vision of Education 4.0. *2019 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC), Engineering, Technology and Innovation (ICE/ITMC), 2019 IEEE International Conference On*. 10.1109/ICE.2019.8792595

Neves, A. R., Costa, J., & Reis, J. (2021). Using a Systematic Literature Review to Build a Framework for University-Industry Linkages using Open Innovation. *Procedia Computer Science*, *181*, 23–33. doi:10.1016/j.procs.2021.01.095

O'Neil, J. M., Newton, R. J., Bone, E. K., Birney, L. B., Green, A. E., Merrick, B., Goodwin-Segal, T., Moore, G., & Fraioli, A. (2020). Using urban harbors for experiential, environmental literacy: Case studies of New York and the Chesapeake Bay. *Regional Studies in Marine Science*, *33*, 100886. Advance online publication. doi:10.1016/j.rsma.2019.100886

Oke, A., & Fernandes, F. A. P. (2020). Innovations in Teaching and Learning: Exploring the Perceptions of the Education Sector on the 4th Industrial Revolution (4IR). *Journal of Open Innovation*, *6*(2), 31. Advance online publication. doi:10.3390/joitmc6020031

Pallot, M., & Pawar, K. (2012). A holistic model of user experience for living lab experiential design. In *2012 18th International ICE Conference on Engineering, Technology and Innovation* (pp. 1–15). IEEE. 10.1109/ICE.2012.6297648

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pan, Y., Oksavik, A., & Hildre, H. P. (2020). *Simulator as a Tool for the Future Maritime Education and Research: A Discussion*. Academic Press.

Ramírez, M.-S., & García-Peñalvo, F.-J. (2018). Co-Creation and Open Innovation: Systematic Literature Review. *Comunicar: Media Education Research Journal*, *26*(54), 9–18. doi:10.3916/C54-2018-01

Schuurman, D., De Marez, L., & Ballon, P. (2015). Living Labs: a systematic literature review. Open Living Lab Days 2015.

Schwab, K. (2017). *The fourth industrial revolution*. Currency.

Schwartz, L. (2019). *How to Use Case Studies in Your Classroom to Make Learning Relevant*. Edutopia.

Song, I.-Y., & Zhu, Y. (2017). Big Data and Data Science: Opportunities and Challenges of iSchools. *Journal of Data and Information Science*, *2*(3), 1–18. doi:10.1515/jdis-2017-0011

Urbinati, A., Chiaroni, D., Chiesa, V., & Frattini, F. (2020). The role of digital technologies in open innovation processes: An exploratory multiple case study analysis. *R & D Management*, *50*(1), 136–160. doi:10.1111/radm.12313

van Geenhuizen, M. (2018). A framework for the evaluation of living labs as boundary spanners in innovation. *Environment and Planning C. Politics and Space*, *36*(7), 1280–1298. doi:10.1177/2399654417753623

Vélez-Rolón, A. M., Méndez-Pinzón, M., & Acevedo, O. L. (2020). Open Innovation Community for University-Industry Knowledge Transfer: A Colombian Case. *Journal of Open Innovation*, *68*(4), 1–17. doi:10.3390/joitmc6040181

Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.

KEY TERMS AND DEFINITIONS

Bulk Carrier: The large single-deck ship, say > 10,000 dwt, which carries unpackaged cargo.

EEA Grants: Through the European Economic Area (EEA) Agreement, Iceland, Liechtenstein, and Norway are partners in the internal market with the Member States of the European Union. To promote a continuous and balanced strengthening of economic and trade relations, the parties to the EEA Agreement have established a multi-annual Financial Mechanism known as the EEA Grants. The EEA Grants established the objective of reducing social and economic disparities in Europe by strengthening bilateral relations with the Beneficiary States.

ENIDH (Escola Superior Náutica Infante D. Henrique): The public polytechnic higher education institute in Portugal and the only national college dedicated to the training of Merchant Marine Officers and senior managers of the Maritime-port sector in the areas of Intermodality, Logistics, and Management.

Growth Mindset: The term used to describe people's beliefs about learning and intelligence means more than just accepting feedback and being open-minded. It is about taking feedback, learning from experience, and coming up with strategies for improvement.

Knowledge Transfer: The communication activities of knowledge between projects and organizations are affected by the other's expertise when at least one of them is affected.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Living Labs: The methodology of innovation that enables collaborative learning by users, producers, and researchers in a real-life environment, in which user needs are central.

Open Innovation: The new business management model for innovation that promotes collaboration with people and organizations outside the company, including high education institutions, tech centers, and other knowledge sources.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 25
The Contribution of Data Science Applied to Customer Relationship Management:
A Systematic Literature Review

Dora Maria Simões
https://orcid.org/0000-0002-9380-4475
University of Aveiro, Portugal

ABSTRACT

In the face the contemporary world lives, and the consequent data produced at an unprecedented speed through digital media platforms, the data are nowadays called the new global currency. It raises numerous opportunities to improve outcomes in businesses, namely at the level of customer relationship management (CRM) strategies and their systems. Nevertheless, how analytics can be applied and support the customer relationship processes seems unclear for academics and industries. To better connect customer relationship processes needs and what data science analytics can offer, this chapter presents a systematic literature review around the concepts, tools, and techniques behind this field, looking particularly on customer acquisition and customer retention in businesses. The outcomes highlight that academic researcher works in this field are very scare and recent. Searching the Scopus and Web of Science databases resulted in only 12 documents from 2013 to 2020, eight of them published in the last two years.

INTRODUCTION

In face the contemporary world lives, and the consequent data produced at an unprecedented speed through digital media platforms, the data are nowadays called the new global currency. It rises numerous opportunities to improve outcomes in businesses, namely at level of competitive advantages and support of decision-making processes (X. Wang, Nguyen, & Nguyen, 2020; C. H. Wang & Lien, 2019; Waller & Fawcett, 2013). Customer relationship management (CRM) is an emergent business and marketing strategy that aims to create and maintain profitable customer relationships by designing and delivering

DOI: 10.4018/978-1-7998-6985-6.ch025

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

superior value propositions. It is based on high-quality customer-related data and leveraged by digital information and communication technology (Buttle & Maklan, 2019; Laudon & Laudon, 2020). Data science. When customer relationship management (CRM) intersects with data science, uncountable opportunities for business emerge. Data science and big data, more specifically network analysis, social media, sentiment analysis, text mining, and information diffusion are application focus of analytics, also, on marketing and business (Camacho et al., 2020). In summary sense, data science can be defined as a multi-disciplinary field that uses scientific methods, techniques, and algorithms, to extract useful knowledge from structured and unstructured data. Big data refers to the use of different methods and techniques to analyze, and systematically extract information from data sets that are too large, or complex, to be dealt with traditional data-processing algorithms. Nevertheless, how analytics can be applied to customer relationship processes is still unclear as the scarce academic publications (De Caigny et al., 2020; Yue, 2020; Sung, Zhang, Higgins, & Choe, 2016) and known industry cases proved. These few developments concerning the state of the art focused on cited concepts is inconsistent with the remarkable advances in artificial intelligence and internet of things in the last decade. To better connect customer relationship processes needs and what data science analytics can offer, this chapter presents a systematic literature review around the concepts, tools and techniques behind the increasing field of data science applied to CRM processes, looking particularly on customer acquisition and customer retention in businesses (X. Wang, Nguyen, & Nguyen, 2020; Iwashita, 2019; C. H. Wang & Lien, 2019; Semrl & Matei, 2017).

The present chapter aims to analyze the literature published in main academic databases - Scopus (by Elsevier) and Web of Science (by Clarivate), under the umbrella of "data science" and "customer relationship management" key terms. The main goal is to fill the gap looking to the published scientific works and identifying tendencies to anticipate the future. The results present an overview of the most relevant themes exploited, their applications, the followed methodologies, removing the veil to other areas to explore.

The chapter is organized as follows: after this introduction the report of the methodologic process of the systematic literature review was performed, presenting the flowchart of the main phases and its steps. Then the results by each phase (input, processing, and output) are reported. The output concerns descriptive and thematic analyses. Finally, the future research directions and a conclusion of the study are highlighted.

METHODOLOGY

One way to achieve greater rigor and better levels of reliability in a literature review is to adopt a systematic approach. In the beginning of the present centenary, Tranfield and colleagues (2003) proposed that systematic literature reviews should be used to develop decision-making evidence to overcome unsystematic processes of literature review and to identify key areas to research. Afterwards, Brereton and colleagues (2007) added that a systematic literature review allows the researcher to make a rigorous and reliable assessment of the research carried out within a specific topic. It is an instrument to map works published on the specific research topic so that the researcher is able to elaborate a synthesis of the existing knowledge on the subject (Levy & Ellis, 2006). The result must be the "state of the art" and demonstrate that the research in question contributes with something new to the existing body of knowledge. Supported mainly in works of Tranfield, Denyer, and Smart (2003) and Levy and Ellis (2006), but also of Moher, Liberati, and Tetzlaff (2015) and Sampaio and Mancini (2007), the methodological

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

approach of this research is mainly supported in three phases (six steps): input; processing and output (see *Figure 1*). In "input" phase there are the data sources and the techniques used during analysis - "processing" – and then, the "output" is generated (synthesis report).

The input phase begins (step one) with the definition and presentation of the main goal of this research: "Determine the most recent developments and applications of data science techniques in customer relationship management (CRM) contexts". After that, the step two continues with the process of data sources identification requiring the definition of rigorous string that suits in the bibliographic databases selected.

The processing phase begins exporting the documents and their references to the Mendeley Desktop – version 1.19.4 - where it was made a preliminary relevance analysis. Consequently, the relevant data about each publication was also exported to Microsoft Excel 2016 and organized in a tabular form. The records were sequentially numbered and sorted by author name, year of publication, title of publication, author keywords, source, publisher, document type, bibliographic database, and number of citations. Based on this, it was possible produce a descriptive analysis, presented by circular and line graphs, showing number of publications by year and distribution of publications by document type. Also, the sources and their respective publishers are analysed in perspective of the influence of their Impact Factor on citations of the papers. After that, and with the goal of produce a thematic analysis, the content of each publication is examined. The goal is to make a critical analysis canvass of the content of the papers, concerning, in each study, the field of action, the methodological approach, and the analytical platforms (step four). To complete this thematic (content) analysis a word frequency query and a cluster analysis by words similarity using NVIVO – version release 1.3 (535) are performed. The approach to process the thematic analysis on step 4 is presented in *Figure 2*.

The research follows to the output phase with a well-structured synthesis of the analyses made in the precedent phase - step five. To support this, a co-word analysis using VOSviewer - version 1.6.16 was performed. Finally, a conclusion about the evidence is presented enhancing the value proposition of the research outcomes to the academy and to the industry (step 6).

Through the application of systematic literature review, this work in addition to provide an overview of the current research being done in the domain of data science technologies and practices applied in customer relationship management strategies, also intends to be helpful for researchers and industries looking for a potential on the subject, by presenting ideas for applying data science to enhance new and innovative CRM strategies.

SYSTEMATIC LITERATURE REVIEW

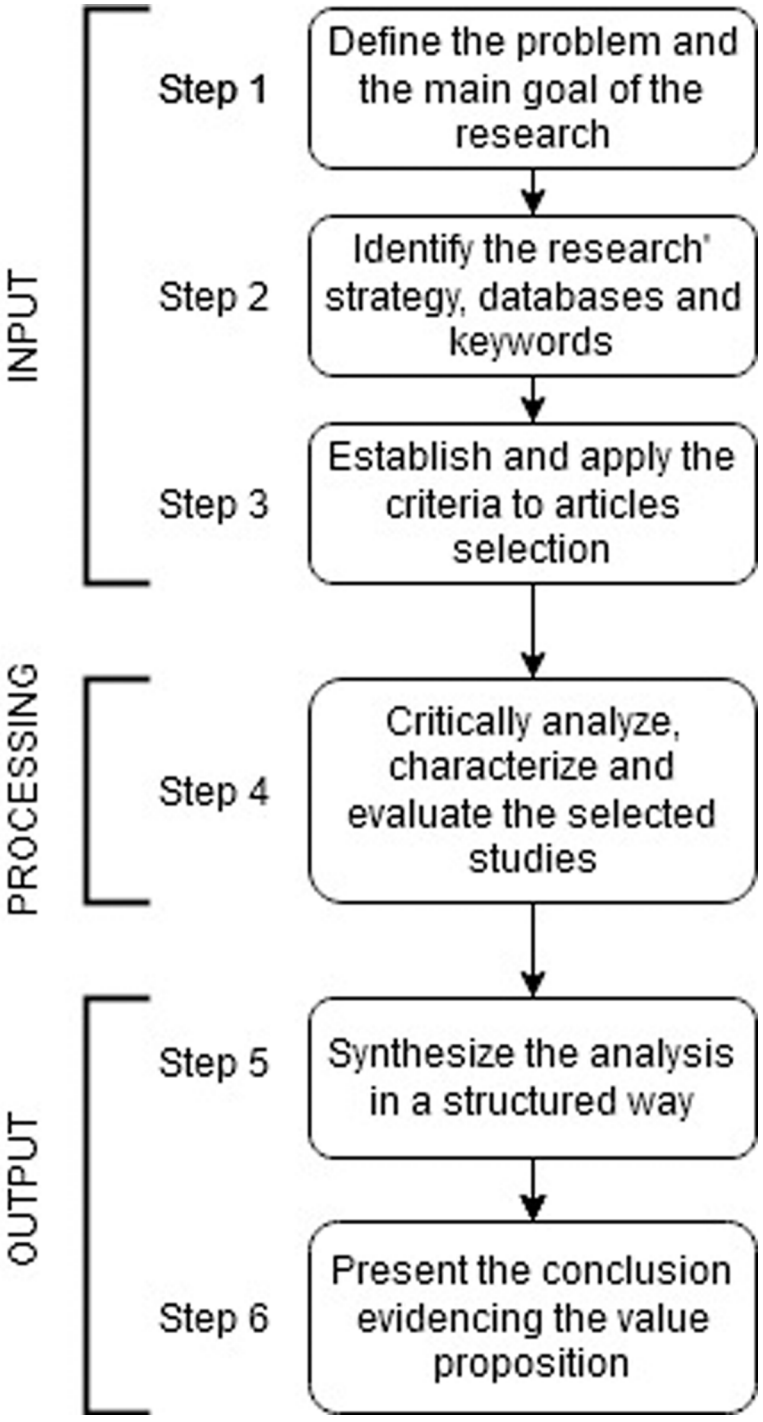
The search was performed on September 29, 2020 following the methodologic procedure described in the previous section. Here it is reported the principal results in a structured way. First, it is presented the strings and the publications output. Second, the guidelines and the outcomes of the descriptive and thematic analyses. And, in third order, it is reported a critical synthesis of the global literature review ran.

Input Steps

In this study are used the Scopus (by Elsevier) and Web of Science (by Clarivate) databases, restricting the results to articles and proceeding papers as document types. Concerning the goal of identify the publications related to research works around the application of data science in customer relationship

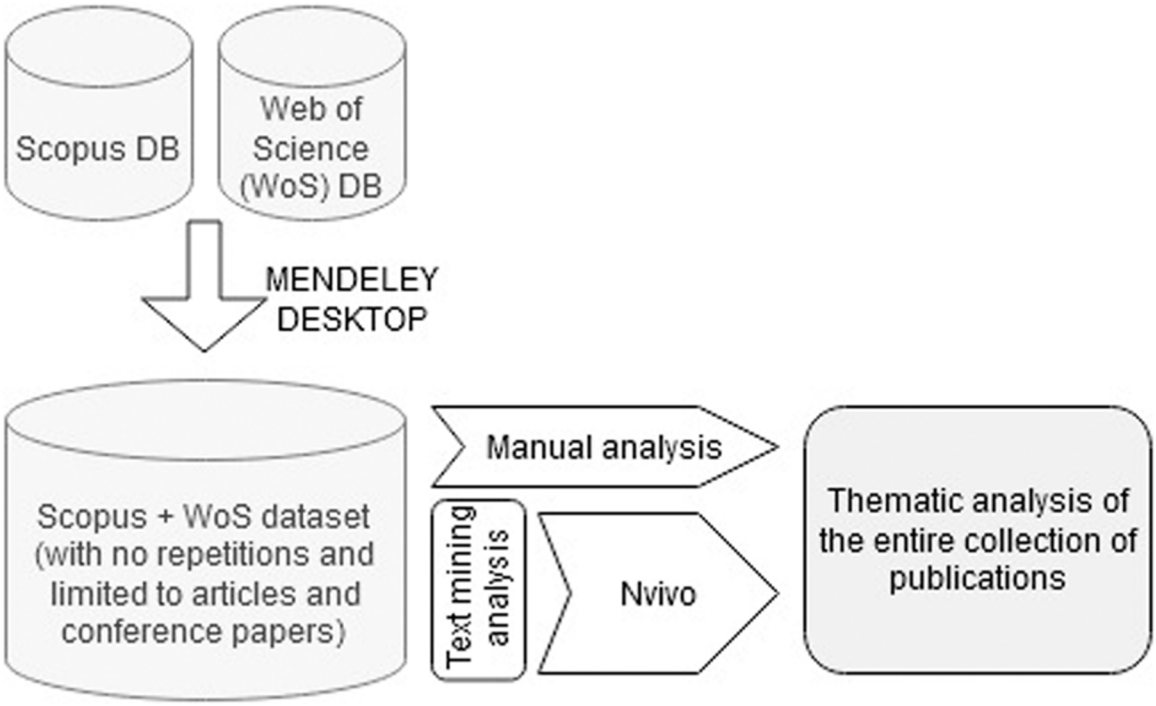
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Methodological approach' flowchart of systematic literature review (adapted from Levy & Ellis, 2006 and Sampaio & Mancini, 2007)



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Thematic analysis approach

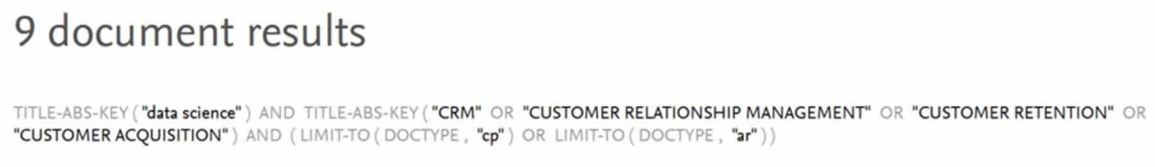


management, and whenever possible, with focus on customer acquisition or retention processes, it was defined to use the following key terms in the search: "data science" and "customer relationship management" or "CRM" (due sometimes it is used the acronym) or the terms "customer acquisition" or "customer retention", because they are terms related to processes strongly closed to CRM strategy. So, attending to these considerations and that these should be founded in the title, the abstract and/or the keywords defined by the authors in the published documents, in the first database, it was used the String 1 (*Figure 2*): TITLE-ABS-KEY ("data science") AND TITLE-ABS-KEY ("CRM" OR "CUSTOMER RELATIONSHIP MANAGEMENT" OR "CUSTOMER RETENTION" OR "CUSTOMER ACQUISITION") AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar")), being obtained 9 documents. In the second database, it was used a similar string (adapted to the database' search codification) – String 2 (*Figure 3*): TS= ("data science") AND TS= ("CRM" OR "CUSTOMER RELATIONSHIP MANAGEMENT" OR "CUSTOMER RETENTION" OR "CUSTOMER ACQUISITION")refined by document types: ARTICLE OR PROCEEDINGS PAPER - obtaining 8 documents, 1 of which is repeated and 4 of them being common to SCOPUS results.

So, applying exclusion and inclusion criteria cited (step three), the total number of the documents is 12. In this case, as all documents are about recent articles, published between 2013 and 2020, in English language, the unique limitation applied was consider articles (ar) and conference paper (cp) as showed in the strings presented above. In the string of Web of Science Core Collection, the term TS means the same as TITLE-ABS-KEY in Scopus, meaning that search is made in title, abstract and keywords of the publications.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

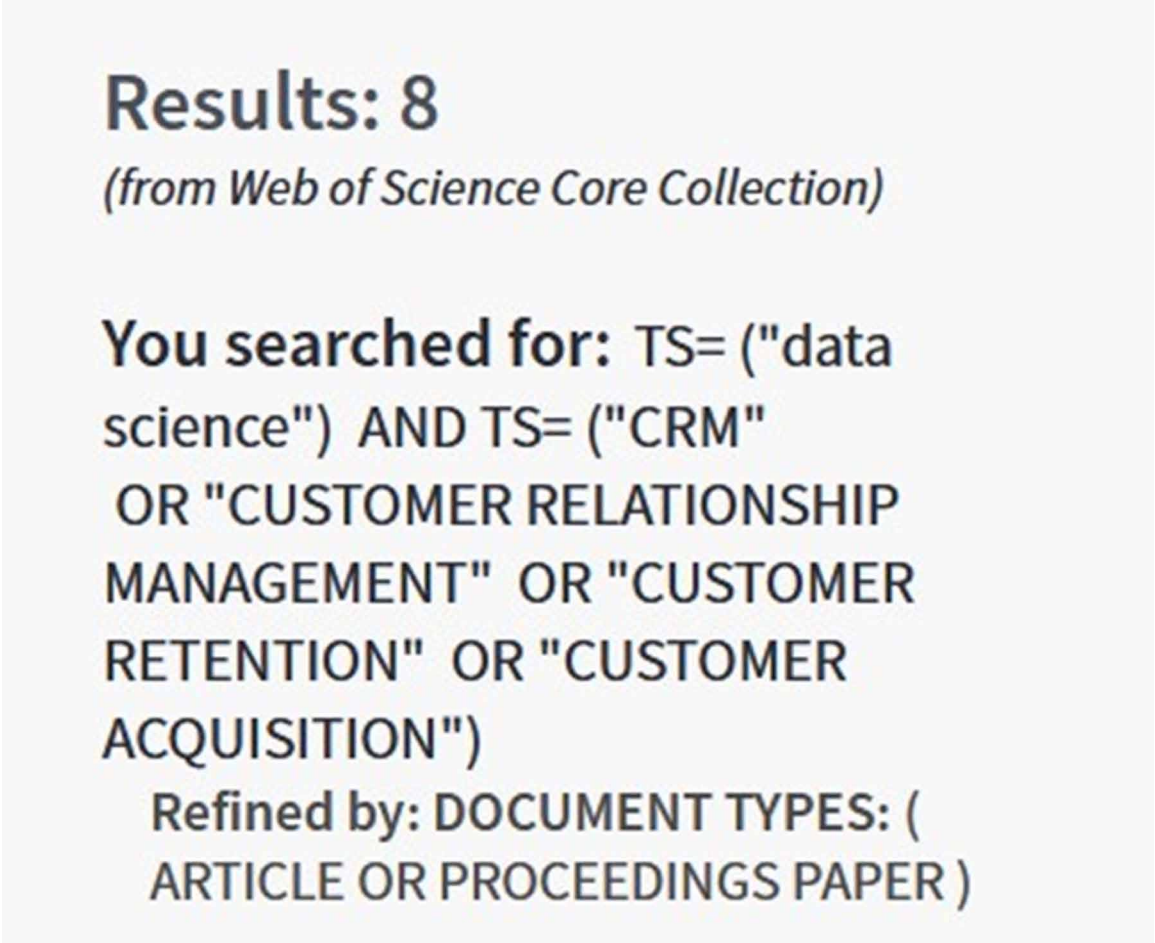
Figure 3. String and results of Scopus



Processing Step

As explained previous, the processing phase involve two types of analyses: descriptive and thematic (content). In the descriptive analysis, the goal was to make a critical analysis under the tendencies of the number of publications per year and its distribution per document type (article or conference paper). Also, it is canvassing the quality of the papers, concerning the sources of publication (journal or conference) and respective publishers, the impact factor and possible correlation with the citation numbers.

Figure 4. String and results of Web of Science



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

The thematic analysis intents to improve the knowledge on each study, namely, the field of action, the methodological approach followed, and the analytical platforms used. Also, a complementary content analysis is performed following the methods available through NVIVO – software provided by QSR International.

Descriptive analysis

After exporting the documents and their references to the Mendeley Desktop software – a bibliographic manager software – the relevant data about each publication was configured in a tabular form, with support in Microsoft Excel 2016 software. In *Table 1*, it is possible to see the result of this, with the records sequentially numbered and sorted by author name, year of publication, title of publication, source, publisher, document type, bibliographic database, and number of citations.

Analyzing the data in *Table 1* it was possible determine the number of publications per year (*Figure 5*) and the frequency' distribution of publications per document type (*Figure 6*). Looking the tables, it is possible to see that all publications are very recent (in a time interval of minus of one decade) and only in last year (2019) and presently (2020) we are assisting a some more activity, with an increase up to 4 publications in each year. Before (till 2018) it was registered no more than 1 publication by year and 0 in the years of 2015 and 2018. The publications are distributed equally by article and conference paper types.

The sources and their respective publishers had only one publication in the timespan of the study (2013-2020). Comparing the Impact Factor (Academic Accelerator, 2021) of each one with the number of citations of the papers, it was determined, and despite of their recency, that the ones published in journals with higher impact factor are the most cited (see *Table 2*). In the *Table 1* it is showed that they are the paper 1 (cited 3 times) - published in 2020 by Elsevier in Journal of Decision Support Systems – and the paper 9 (cited 95 times) – published in 2013 by Council of Supply Chain Management Professionals in Journal of Business Logistics. The other documents in analyze in this study with citations (both cited two times), till the moment, are the papers 7 and 8 – published by renowned Institute of Electrical and Electronics Engineers Inc. (IEEE) in International Conference on Behavioral, Economic, and Socio-Cultural Computing and International Conference on Data Science and Advanced Analytics, respectively, in 2017 and 2016.

Surprisingly, concerning their impact factor, they are the Journal of Industrial and Production Engineering and Lecture Notes in Computer Science published by Springer and Taylor and Francis Ltd., respectively, with no citation in their papers (2 and 10). One reason could be their recency due they were published in last year (2019). Other papers that it was expected more visibility are the 5 and the 11, both published by the Association for Computing Machinery (ACM), through the proceedings in conferences held in 2014 and 2020.

The analysis of the authors looks through that there aren't assiduous authors, in the context of this study. In the sample, each author or co-author participate only one time in each paper noticed. Also, there are no evidence of social relationship between authors through different publications.

Thematic analysis

With the goal of produce a thematic analysis, the content of each publication was examined manually. The thematic (content) analysis performed was firstly oriented by the keywords, the abstract and the title, and then all the content was analysed (reading each publication) to better understand their focus. With

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. List of publications sorted by authors' name

Nº	Authors	Year	Title	Source	Publisher	Citations
1	(De Caigny, Coussement, & De Bock, 2020) (1) (2) (3)	2020	Leveraging fine-grained transaction data for customer life event predictions	Journal of Decision Support Systems	Elsevier B.V.	3/0
2	(de Castro Neto, Julia, Paiva, Carvalho, Junior, Peres, et al., 2019) (2) (4)	2019	Improving the AHT in telecommunication companies by automatic modeling of call center service	Lecture Notes in Computer Science	Springer Verlag	0
3	(Delgosha, Hajiheydari, & Saadeatmantesh, 2020) (1) (3)	2020	Semantic structures of business analytics research: applying text mining methods	Information Research - An International Electronic Journal	Univ. Sheffield - Dept. Information Studies	0
4	(Iwashita, 2019) (1) (3)	2019	Transitional method for identifying improvements in video distribution services	International Journal of Networked and Distributed Computing	Atlantis Press Sarl	0
5	(Jussila et al., 2014) (2) (4)	2014	New era of business analytics - making sense of business ecosystems	International Academic MindTrek Conference	Association for Computing Machinery, Inc (ACM)	0
6	(Salas-Rueda & Salas-Silis, 2019) (1) (3)	2019	Wevideo: cloud service useful for students during the construction and presentation of audiovisual contents?	Vivat Academia	Univ. Complutense Madrid, Servicio Publicaciones	0
7	(Semrl & Matei, 2017) (1) (2) (4)	2017	Churn prediction model for effective gym customer retention	Conference on Behavioral, Economic, and Socio-Cultural Computing	Institute of Electrical and Electronics Engineers Inc. (IEEE)	2/0
8	(Sung, Zhang, Higgins, & Choe, 2016) (1) (2) (4)	2016	Data-driven sales leads prediction for everything-as-a-service in the cloud	International Conference on Data Science and Advanced Analytics	Institute of Electrical and Electronics Engineers Inc. (IEEE)	2/2
9	(Waller & Fawcett, 2013) (2) (3)	2013	Click here for a data scientist: big data, predictive analytics, and theory development in the era of a maker movement supply chain	Journal of Business Logistics	Council of Supply Chain Management Professionals	95
10	(C. H. Wang & Lien, 2019) (1) (2) (3)	2019	Combining design science with data analytics to forecast user intention to adopt customer relationship management systems	Journal of Industrial and Production Engineering	Taylor and Francis Ltd.	0/0
11	(X. Wang, Nguyen, & Nguyen, 2020) (2) (4)	2020	Churn prediction using ensemble learning	International Conference Proceeding Series	Association for Computing Machinery (ACM)	0
12	(Yue, 2020) (2) (4)	2020	Topological data analysis of two cases: text classification and business customer relationship management	Journal of Physics: Conference Series	Institute of Physics Publishing	0

(1)Web of Science/CLARIVATE
(2)Scopus/ELSEVIER
(3)Article
(4)Conference paper

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Publications per year

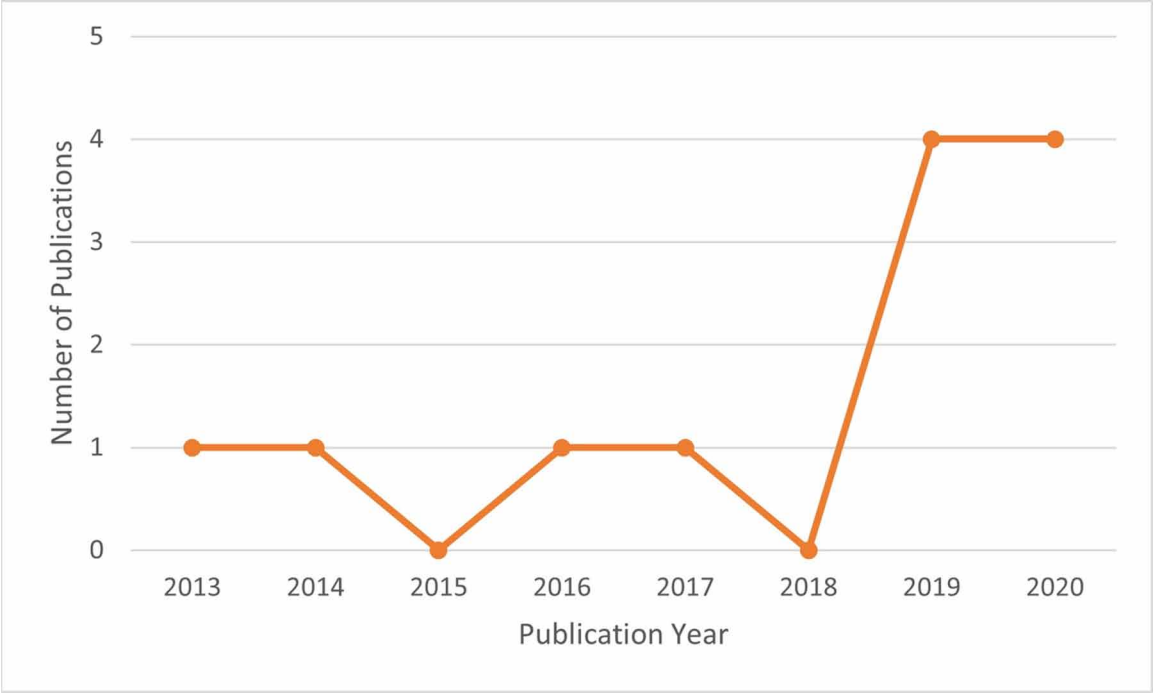
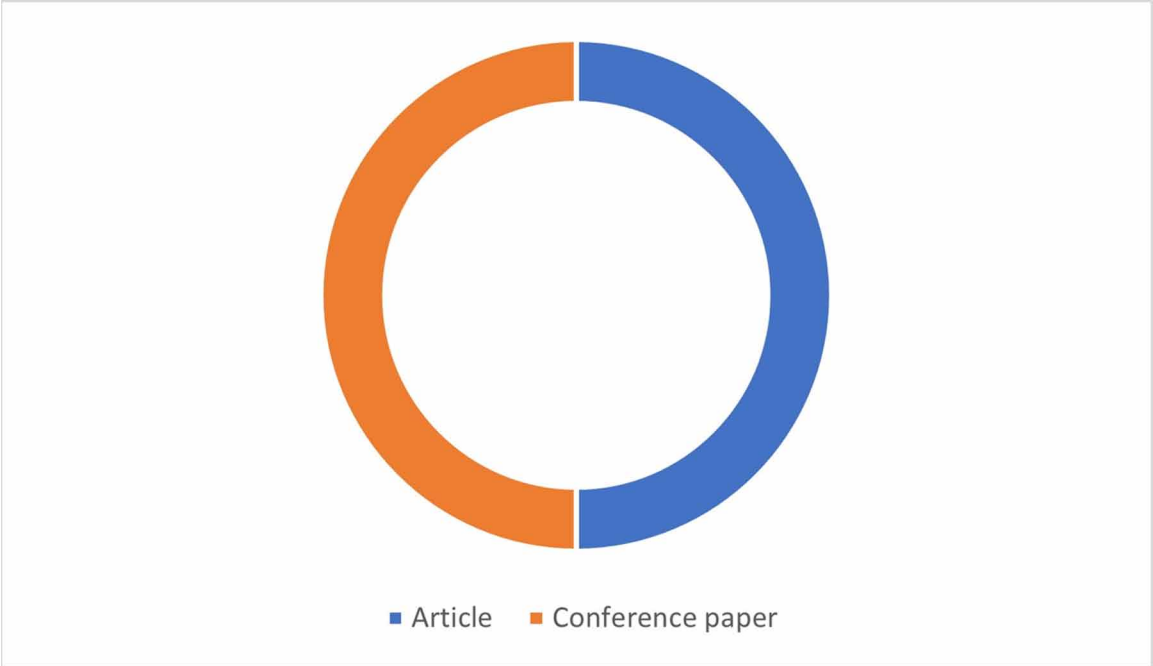


Figure 6. Publications per document type



EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Impact factor of publication' sources

Source	Publisher	Impact Factor
Journal Decision Support Systems	Elsevier	5.420
Journal of Business Logistics	Council of Supply Chain Management Professionals	3.850
Journal of Industrial and Production Engineering	Taylor and Francis Ltd.	1.960
Lecture Notes in Computer Science	Springer	1.170
Information Research - An International Electronic Journal	Univ. Sheffield - Dept. Information Studies	0.763
Journal of Physics: Conference Series	Institute of Physics Publishing	0.540
International Journal of Networked and Distributed Computing	Atlantis Press Sarl	0.200
International Conference Proceeding Series	Association for Computing Machinery (ACM)	-
International Conference on Data Science and Advanced Analytics	Institute of Electrical and Electronics Engineers Inc. (IEEE)	-
International Academic MindTrek Conference	Association for Computing Machinery (ACM)	-
International Conference on Behavioral, Economic, and Socio-Cultural Computing	Institute of Electrical and Electronics Engineers Inc. (IEEE)	-
Vivat Academia	Univ. Complutense Madrid, Servicio Publicaciones	-

focus on these it was intended identify the key terms, the field of action, the methodological approach, and the analytical platforms (see *Table 3*).

Based on the results of content analysis presented above in *Table 3* and in the analysis of all corpus of publications clustered by word similarity, using Pearson Correlation Coefficient through NVIVO software, a summary of each set of publications with more similarity is described. The cluster analysis generated a diagram that cluster the selected files (the 12 publications) together if they have many words in common. The results can be viewed as a dendrogram or cluster map (see *Figure 7*). The summary proceeds in ascending order of relevance to focus of this chapter, that is also according the cluster tree map.

The paper #6 (Salas-Rueda & Salas-Silis, 2019) differs somewhat from the main objective of this systematic literature review because its focus is on analysis of the impact of the Wevideo cloud service during the teaching-learning process in higher education, and where the key term CRM appear resulting from the theme explored by the students in their practical exercises. However, this paper refers the use of RapidMiner – and relevant analytical platform for mining data. Also, the paper #9 (Waller & Fawcett, 2013) is irrelevant for this study because it is an editorial paper that proposes a 2x2 model to explain the role of predictive analytics in the theory development process and incentive researchers to publish in the journal, contributing to explaining the role of predictive analytics in the theory-building process.

Concerning the following set of papers: #3 from Delgosha and colleagues (2020) and #5 from Jussila and colleagues (2014), the first one - paper #3 - presents a study where the main research goal was identifying the major research topics and trends using text mining techniques. The authors begin distinguishing the three main hierarchical categories: descriptive, predictive, and prescriptive analytics (Delen, 2014). Descriptive analytics focusing on 'what happened in the past and why', while predictive analytics using models to answer the question of 'what will happen in future?', and prescriptive analytics going further to provide guidance by evaluating possible scenarios and answering, 'what course of

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

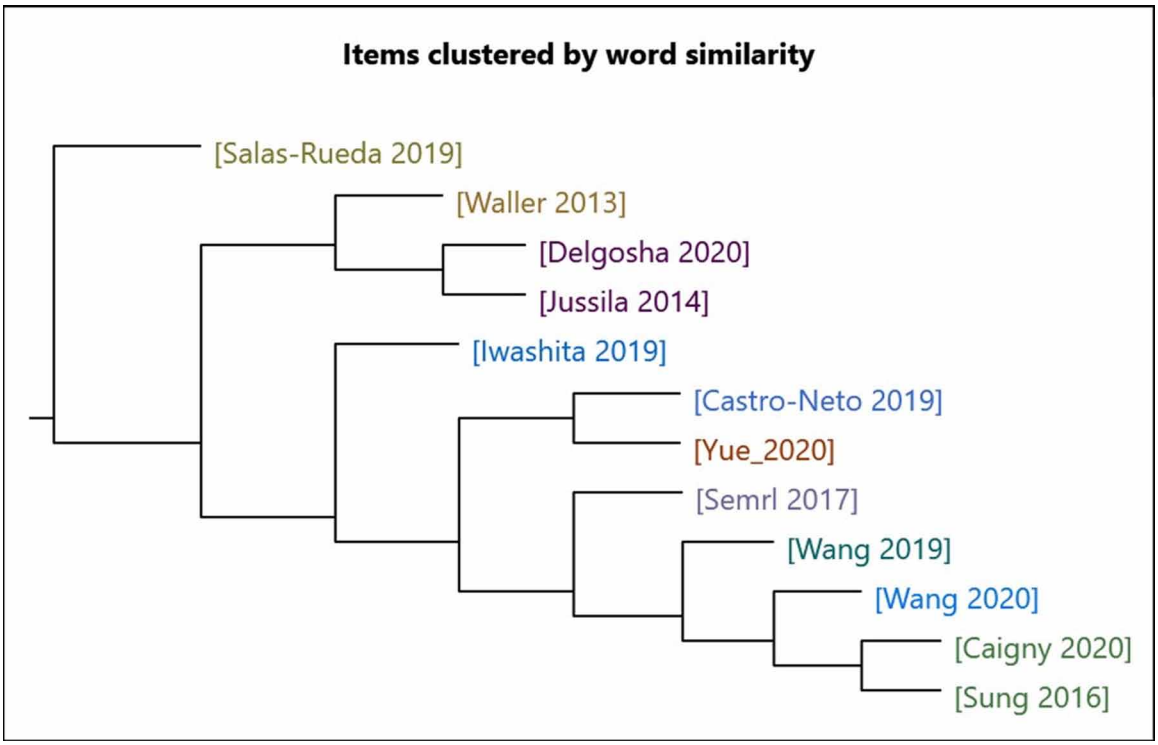
Table 3. Thematic analysis of publications

Nº	Key terms[1]	Field of action	Methodologic approach	Analytical platforms
1	data science, customer relationship management, CRM	Validation of a new customer life event prediction model combining aggregated customer data (demographics, behavior, and contact) with fine-grained transaction data (that preserve recency, frequency, and monetary value information of the transactions) of a financial services industry, to predict four different customer life events.	Customer life event prediction (CLEP) method. Area Under the receiver operating characteristics Curve (AUC), Top Decile Lift (TDL), Wilcoxon signed-ranks tests and Wald statistics.	-
2	data science, customer relationship management, CRM	Enhance the customer service of an ISP company through performing a modelling of its CRM data warehouse and using such model to improve the call center scripts, so as to reduce the Average Handle Time (AHT).	Data Mining to induce classification rules able to predict the need for technical visit considering the customer problems.	-
3	data science, customer retention	Identify major research topics and trends using text mining techniques based on Web of science and Scopus databases.	Two text mining methods: co-word analysis and topic modelling.	VOSviewer BigML
4	data science, customer retention	Describes issues related to using customer satisfaction (CS) concept as a key indicator to customer retention, and its application to video distribution services (like Youtube, Netflix, etc), following a transitional evaluation approach.	Three evaluation methods for finding improving factors. Questionnaire data analysis oriented for the need of transitional approach.	-
5	data science, CRM	Discuss the business analytic applications and respective methods and industries.	Machine Learning methods, social network analysis and information visualization methods, and simulation methods.	NoSQL databases Hadoop
6	data science, CRM	Analyze of the impact of the Wevideo cloud service during the teaching-learning process in higher education. Using RapidMiner it is identified 4 predictive models.	Quantitative and qualitative research, predictive models by means of decision tree technique.	Wevideo cloud service RapidMiner
7	data science, customer retention	Predicting customer churn from contractual and behavioral perspectives, applied to fitness industry.	Machine Learning research.	AzureML BigML
8	data science, customer relationship management	Proposal of a data-driven iterative sales lead prediction framework for cloud everything as a service (cloud-based SaaS service), with the objective of to help the business team to nurture valuable potential customers, classifying (grouping) by paying and non-paying users.	RFDL (Recency, Frequency, Duration and Lifetime) analysis method, an extension of the RFM (Recency, Frequency, and Monetary value) model, Jenks natural breaks algorithm, and classification data mining techniques. BizDevOps process and DevOps deliverables.	IBM cloud platform (cloud-based relational data warehouse)
9	data science, customer relationship management	Editorial paper that proposes a 2x2 model to explain the role of predictive analytics in the theory development process and calls for other contributions in this domain.	2x2 model (matrix of explanation and prediction) pointing research questions relevant to practitioners.	-
10	data science, customer relationship management, CRM	(1) identifying the causalities between design features and behavioral intention, (2) deriving the priorities of design features, and (3) forecasting user intention to adopt CRM systems.	Behaviour-science (user perceptions and Behaviour intentions): Technology-Acceptance-Model (TAM) and design-science: Design-Based Research (DBR). Classification & Regression Tree (CART), Random Forest (RF) and Adaptive Boosting (AdaBoost).	-
11	data science, customer retention	A comparative study of the most widely used classification methods on the problem of customer churning in the telecommunication sector, using a public dataset from the Kaggle competition.	Machine learning classification models/techniques supported in three phases method: data pre-processing, model exploration, and fine-tuning the system. Classifiers: LightGBM, XGBoost, Decision Tree (DT) and Random Forest (RF).	-
12	data science, customer relationship management	Predicting customer loyalty and demand, applied on business-to-business.	Topological Data Analysis (TDA): 1st case - RFM + Prediction method and Time-series clustering technique, 2nd case – Persistent Homology and Mapper algorithms. Time Series Data and Text Classification of Natural Language Processing (NPL)	-

[1]In Title, Keywords or Abstract.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Tree map of publications clustered by word similarity



action should we take?'. They used two text mining methods to analyse content. Co-word analysis was used to calculate relationships among terms based on co-occurrence statistics, and topic modelling was used to identify the main topics based on the latent Dirichlet allocation model. They used to mine and visualise data VOSviewer and BigML, respectively. Latent Dirichlet allocation is a popular technique for content analysis designed to organize massive volumes of documents based on topics (Teh et al., 2007). BigML is a machine learning platform that provides sophisticated tools for statistical natural language processing to identify and model business analytics topics. The second one - paper #5 - discuss the business analytic applications and respective methods and industries. It could be a relevant contribution to deeper the knowledge on big data tools like NoSQL databases and Hadoop.

The paper #4 (Iwashita, 2019) describes issues related to using Customer Satisfaction (CS) concept as a key indicator to customer retention, and its application to video distribution services, following a transitional evaluation approach. The author distinguishes 3 types of video services: (a) video distribution (like Hulu or Netflix), (b) video share (like YouTube or Dailymotion), and (c) video broadcast (like Hikari TV or U-NEXT, or recently, in Portugal, the OPTO from SIC TV). Also, it is considered two types of business models: (1) free-for-viewer and (2) pay-for-viewer. The author explains customer loyalty as generally depending on the relationships among three factors: price, customer expectation (CE), and CS. Price is decided during service optimization. Customer expectations are set based on sales promotions, and the customer experiences satisfaction based on these expectations upon obtaining the goods/services. The good service is considered to be supported by a good strategy when CS is higher than CE. Loyalty decreases when this relationship is reversed, although the damage would be relatively small if CS is greater than price. However, customer retention will be difficult if CS is lower than price. Therefore, CS

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

increases when it is larger than CE. The authors compare three evaluation methods for finding improving factors: evaluation using dissatisfaction scores (method 1), evaluation using difference in satisfaction and dissatisfaction scores (method 2), and transitional evaluation of satisfaction and dissatisfaction scores (method 3). The comparison is proceeded supported in a questionnaire and the data analysis oriented for the need of transitional approach. This is a relevant work highlighting the application of data science in business, and, particularly, in customer relationship management. In this case, specifically with the goal of improve customer satisfaction and consequently customer loyalty (retention).

Analysing now the set composed of the paper #2 (de Castro Neto et al., 2019) and the paper #12 (Yue, 2020), both are focused on Customer Relationship Management (CRM).In the paper #2 it is enhanced the customer service of an ISP company through performing a modelling of its CRM data warehouse and using such model to improve the call center scripts, such as to reduce the Average Handle Time (AHT). The modelling proposed uses data mining to induce classification rules able to predict the need for technical visit considering the customer problems. In the experiments carried out, rules with high predictive accuracy were generated. The results confirm that the generated rules allow for reducing the call AHT and can also be used as a tool for reducing the need for technical visits. Looking in paper #12 and particularly about the first case analyzed (Rivera-Castro et al., 2019), the authors enhance that after constructing a time series, from a qualitative perspective, the benefit of using them is their highly visual component. The marketing analyst can show the results to decisions-makers. From a quantitative side, topological RFM (Recency, Frequency and Monetary value) approach was used as a clustering method to increase the accuracy of a predictive model for loyalty scoring. Another benefit highlighted is the improvement on accuracy of a machine learning model. Thus, they open the door for a CRM predictive pipeline, where the practitioner can segment the customer base, generate personas, do predictions, and communicate to manage both predictions for the personas as well as for the individual users and identify those users diverging from the expected results from their respective personas. Therefore, the provider can obtain a deeper understanding of its customer base to adjust its product offering promotions while being able to generate reliable estimates for their future demand over multiple periods.

The paper #7 (Semrl & Matei, 2017) presents another example of data science application on service of customer relationship management (CRM). In this case, predicting customer churn from contractual and behavioral perspectives, applied to fitness industry. The AzureML and BigML analytical platforms supported the research' work. The last one already cited here in explanation of #paper 3.

The paper #10, from Wang and colleague (C. H. Wang & Lien, 2019), presents a study where the main goal of the authors was (1) identify the causalities between design features and behavioral intention, (2) derive the priorities of design features, and (3) forecast user intention to adopt CRM systems. The study allowed the authors highlighted that the contract-based sectors pay attention to managing call-center monitoring, loyalty programs and churn prevention while non-contract-based sectors focus on the skills of sales force automation, up-selling or cross-selling, and dynamic pricing. Enterprises with prior implementation experiences focus on the construct of perceived usefulness (PU) while the inexperienced enterprises focus on the construct of perceived ease-of-use (PEOU). Expert's users knowledgeable of statistics and database are more interested in mining analytics than the features of exploratory understanding. Also, the study allowed find that the educational training in statistics or database is helpful to improve CRM operational skills for the staff (front-end and the back-end) and the novice users. Yet, forecasting user intention to adopt CRM can achieve better accuracy than forecasting user intention to recommend CRM. The authors enhance to that individual factors are highlighted, only job occupation (front-end vs. back-end) and prior knowledge (expert vs. novice). Following the concept of TAM, only

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

PU and PEOU are treated as the predictors. Other constructs, such as complexity, compatibility, and trialability, can be considered as the mediators or moderators in forecasting user. In opinion of the authors, despite ERP and CRM are popular enterprise applications, CRM is significantly different from ERP in two aspects: (1) ERP handles operational transactions like purchasing, accounting, and inventory control while CRM focuses on managing customer profiles and extending or enhancing customer relationships. (2) ERP needs to be tailored to internal business processes, but CRM depends on timely information to quickly respond to diverse requirements and handle customer complaints. In brief, ERP is more common in manufacturing sectors while CRM is more service-oriented. To address the impacts of organizational characteristics and individual characteristics, a novel TAM that combined design science with data science is presented help CRM developers identify key design features and capture the causation between user perceptions and behavioral intention. The enterprises with prior experiences in the implementation of enterprise applications are more likely to adopt the CRM system than the inexperienced ones. That means, CRM vendors can consolidate customer loyalty by targeting the enterprises who have already installed their ERP before. Meanwhile, for the purpose of market segmentation, CRM vendors should develop customized commercial software to satisfy the diverse requirements of distinct user groups. For individual users, poor knowledge in statistics or database can result in an obstacle in exploiting the strengths of the CRM software. The above-mentioned findings, in turn, suggest both CRM vendors and service companies the necessity of job training and developing different CRM versions (basic and advanced) for satisfying distinct user groups. In particular, the prediction of user recommendation is more difficult than the prediction of user adoption because recommendation needs user satisfaction with CRM systems and/or justification of performance comparisons between " before " adoption and " after " adoption. In the future, the impacts of cloud computing and artificial intelligence on mobile carriers (smartphones or tablets) deserve to be further addressed in designing CRM apps.

The paper #11 (X. Wang, Nguyen, & Nguyen, 2020) presents a comparative study of the most widely used classification methods on the problem of customer churning in the telecommunication sector, using a public dataset from the Kaggle competition. The authors cite the works of De Caigny and colleagues (e.g., De Caigny, Coussement, & De Bock, 2018) - Logit leaf model (LLM) – combination of decision tree and logistic regression, and Coussement and colleagues (Coussement, Lessmann, & Verstraeten, 2017) - improving churn prediction performance by data preparation treatment strategy. In this paper, the authors promote the application of machine learning classification models/techniques supported in three phases method: data pre-processing (including six steps: missing values, outliers, categorical variables, normalization, class imbalance, and feature selection), model exploration, and fine-tuning the system. Also, a top four performing classifiers are identified being the bagging and boosting ensemble techniques (LightGBM, XGBoost), Decision Tree (DT) and Random Forest (RF).

Finalizing this summary analysis about the publications in study, the focus is put on the paper #1 (De Caigny, Coussement, & De Bock, 2020) and on the paper #8 (Sung, Zhang, Higgins, & Choe, 2016). In the first one, the authors defend that managing customer relationships during such life events has crucial strategic importance and can align the company's actions better with future customer behavior. Correct detection of life events enables companies to engage customers with appropriate cross- or up-selling offers and reactively target them with relevant marketing campaigns. In their study, the authors determined that the customer demographics are always the most important for predicting life events despite customer behavior information and can help predict life events with less importance. Yet, the contact between the customer and company have the least importance. With respect to the fine-grained transaction data, the authors note substantial impacts in terms of predicting life events. They represent the most important

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

source of information for moving, birth of a child, new relationships, and end of a relationship. The model that uses the RFM-extended behavioral similarity scores extracts far more predictive information out of the data, which then leads to heightened importance scores. So, they incorporate fine-grained transaction data, and particularly extending the PSN (pseudo-social network) methods with behavioral similarity measures based on RFM variables obtained from payment transactions, significantly improves predictive performance. So, the results show that optimal predictive performance requires a combination of aggregated and fine-grained transactional data. Predicting life events might help to target the right customers with the right promotions, which could give a competitive advantage. In addition to the case study applied to the financial industry, they suggest application of CLEP (customer life event prediction) method to telecommunications or retail industries, meanwhile drawing attention to data quality. Customer life event prediction (CLEP) method offer a framework for integrating fine-grained data in a predictive model (representing the data in a pseudo-social network (PSN) constructed on binary information and continuous data (RFM-based). To measure the predictive performance of the models it was used the area under the receiver operating characteristics curve (AUC) and top decile lift (TDL). To test whether observed differences in the performance measures across the models are statistically significant, the comparison of the classifiers used Wilcoxon signed-ranks tests. The paper #8 (Sung, Zhang, Higgins, & Choe, 2016) propose a data-driven iterative sales lead prediction framework for cloud everything as a service (cloud-based SaaS service), with the objective of to help the business team to nurture valuable potential customers, classifying (grouping) by paying and non-paying users. In this work, the authors propose the RFDL (Recency, Frequency, Duration and Lifetime) analysis method, an extension of the RFM (Recency, Frequency, and Monetary value) model. It is used IBM cloud platform (cloud-based relational data warehouse), BizDevOps process and DevOps deliverables.

As a remarkable point, the word cloud of *Figure 8* obtained through the quality software NVIVO confirm the domain of the publications.

Figure 8. Word cloud of frequency

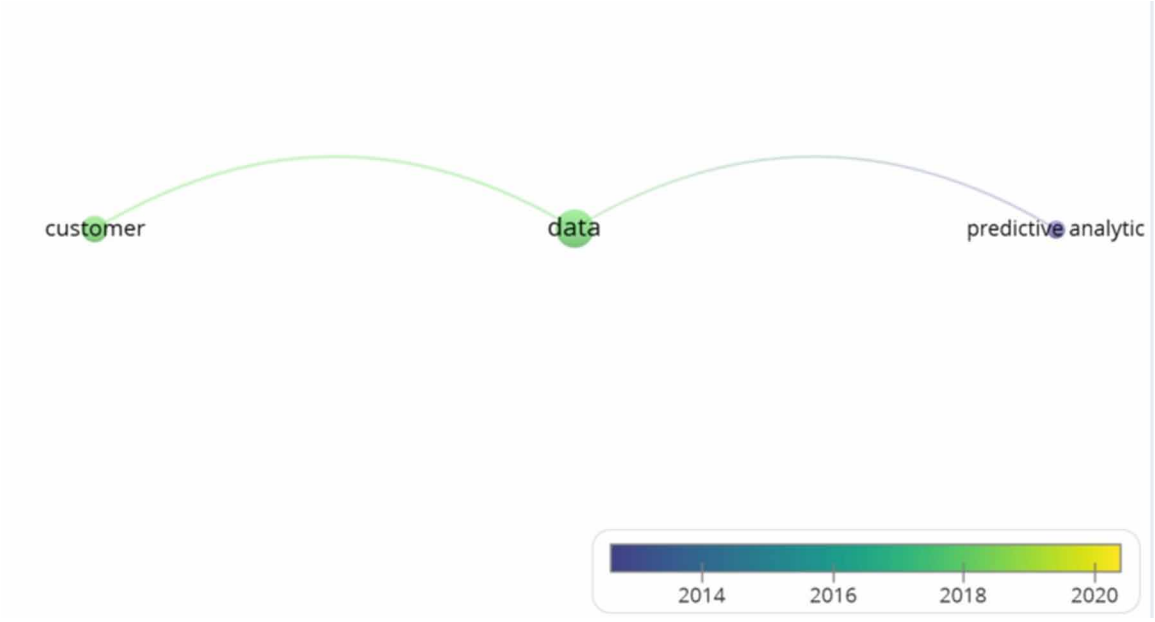
In map (*Figure 8*) it is showed, based on a word frequency query with 100 most frequent, that the main themes of research are around customer and CRM, data analytics and models in business.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Output Steps

The synthesis of the analysis and the value proposition evidence is now presented. It was used VOS-viewer – free software provided by Centre for Science - to mine and visualise scientific texts (Van Eck & Waltman, 2010). The software uses text mining functions to identify related noun phrases, along with a clustering approach to examine the co-occurrence of the words. A co-occurrence map shows that the distance between two terms can be interpreted as an indication of the relatedness of the terms. It applies a distance-based mapping technique referred to as visualisation of similarities (VOS). Co-word analysis is based on two important assumptions: 1) the words used are good representatives for the reflection of its content and 2) the co-occurrence of the two words in various papers indicates their relation so that if the co-occurrence of the keywords is more, their relation is closer (Feng et al., 2017). Through the VOS-viewer software, the *Figure 9* shows a term co-occurrence map based on text data obtained based on the RIS file created from the publications' collection in analysis stored in Mendeley bibliographic software. This represents a co-word analysis network of high-frequency words related to this research, illustrating how common words presented in title and abstract fields of all documents are related to each other.

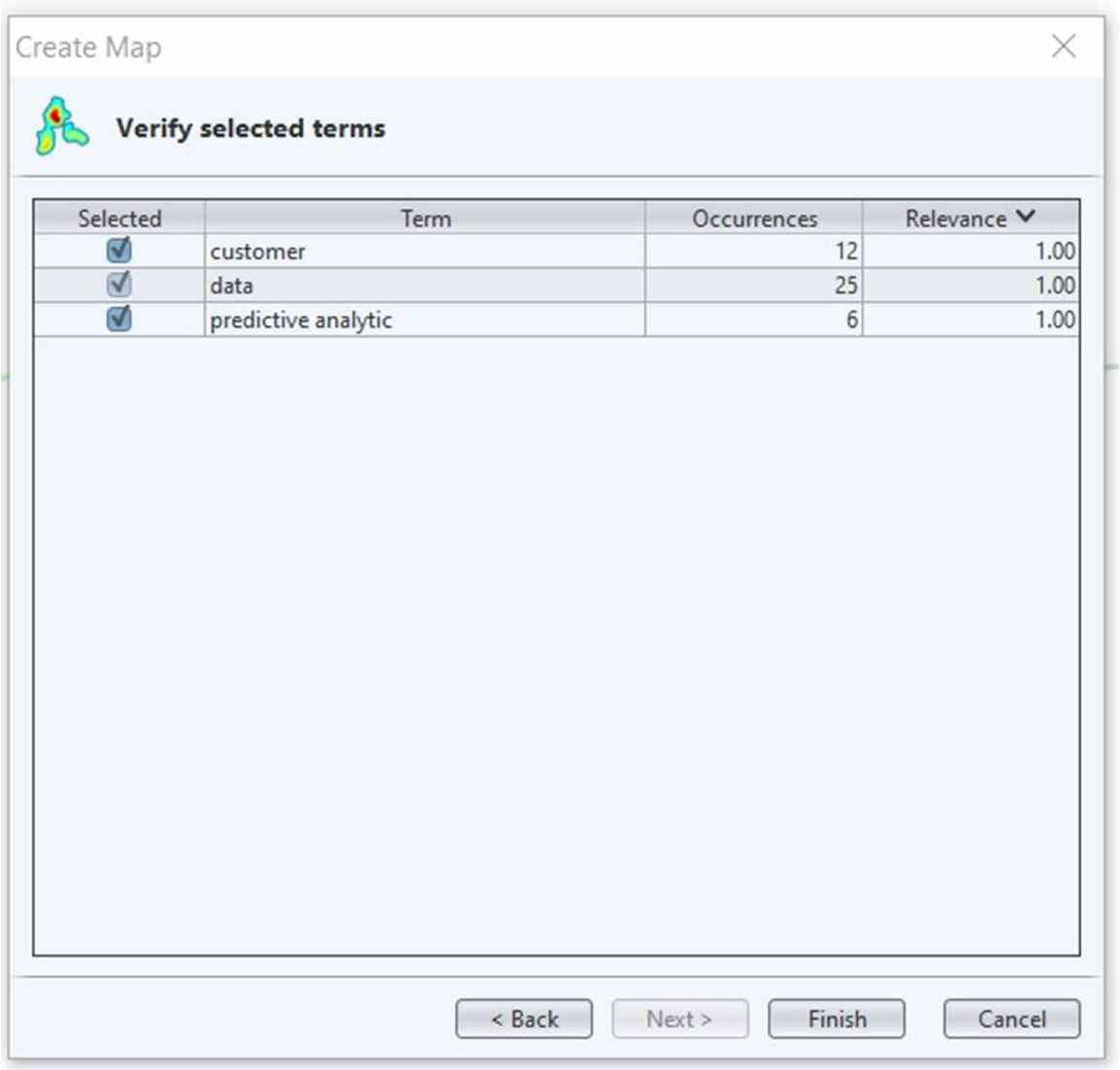
Figure 9. Network and overlay visualization of co-occurrence analysis on titles and abstracts



A full counting method was followed based on a minimum number of occurrences of a term of 6. This network identified 447 terms where 4 meets the threshold. For each of the 4 terms a relevance score was calculated and based on it the most relevant ones were selected based on premise of 60% most relevant. Then, 3 terms were select: customer, data, and predictive analytic (see *Figure 10*).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

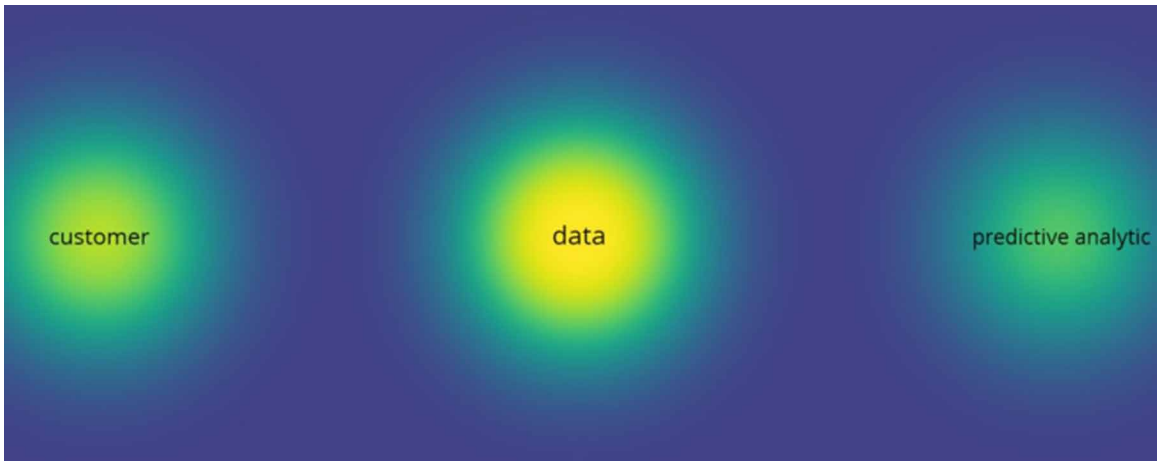
Figure 10. Terms with more occurrences and strength



The three terms were identified with the same relevance - 1.00 – being "data" the most frequent term (25) followed of customer (12) and predictive analytic (6). In this yielded of words, in *Figure 9*: green (data and customer) and violet (predictive analytic). Most important terms are marked with a larger font. Using the 'overlay visualisation' tab in VOSviewer software, the main topics of recent research in data (science) field, the rising and fading trends were identified. So, terms such as data and customer are among the most frequent topics in research. As observed in *Figure 9*, data and customer are two topics that have emerged as trends since mid-2016. Curiously, the term "predictive analytic" was more frequent in the begin of this timespan (2013-2014). More recently (from 2018) is emerging the term "data (science)"- represented with yellow colour in F*igure 9* and *Figure 11* - with a strong incidence but with yet little expression due few publications reference it.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Density visualization of co-occurrence analysis on titles and abstracts



The network, overlay and density visualisation of co-occurrence analysis based on titles and abstracts of publications reinforce the previous explanation that more and more research studies and practical applications works around the data science field applied to customer relationship management strategies and systems will be wellcome.

Truly, this study identified 8 publications centred on core of this research. They are: paper #4 from Iwashita (2019); paper #2 from de Castro Neto et al. (2019); paper #12 from Yue (2020); paper #7 from Semrl and Matei (2017); paper #11 from Xing Wang, Khang Nguyen, and Binh P. Nguyen (2020); paper #1 from De Caigny, Coussement, and De Bock (2020) and the paper #8 from Sung, Zhang, Higgins, and Choe (2016). The paper #10 from Chih-Hsuan Wang and Ching-Yu Lien (2019) despite its focus on CRM systems it is centred on its adoption by the enterprises and their users.

FUTURE RESEARCH DIRECTIONS

Evidence for this study and it was already referred in past in systematic literature review work of Saur-Amaral, Ferreira, and Conde (2013), despite their study have been centered on tourism marketing and consumption theme, the field of relationship marketing was cited as scarcely present, loyalty and customer relationship management tools were identified with few references and no real methodology or model applied to manage the relationship with the tourists and to measure the value of each customer. Eight years passed from the research of these authors, the systematic literature review made and present in this chapter highlights that there are not developments in this field (customer relationship management applied to tourism and its congeners business areas like hotels, travel agencies and companies, and so on). So, our research illustrated that there are clear gaps in the scientific literature and much core to explore in this thematic and in others business areas like telecommunications, bank, education, construction sector, etc.

This research has a few constraints, as it is limited to Scopus and Web of Science databases. The analysis may not be capable of extracting very deep knowledge of the specified context. As the data science applied to customer relationship management knowledge area is in its infancy stage, the keywords used in related studies are not yet standardised. Therefore, for a more complete understanding of this

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

new area, researchers and practitioners might use congeners search terms of data science such as business intelligence, business analy*, big data analy*, data analy* and data mining without caring about the uppercase or lowercase or the spaces within the terms. Thus, they can retrieve documents from various databases as they may be used interchangeably by this search terms and "customer relationship management" or "CRM" and other terms. Future studies can use complementary bibliometric techniques (e.g., co-citation analysis, co-author analysis and bibliographic coupling) to give special attention to authors who provide links between the major knowledge groups would also be of interest, such as the kinds of work these links represent and the questions they address.

CONCLUSION

The aim of this manually and computational literature analysis was to identify the knowledge structure, review the main and influential trends and reveal hidden themes in the field of data science applied to customer relationship management in business. Text mining methods were used to analyse published documents in the Scopus and Web of Science databases. This study investigated the semantic structure analysing 12 related research documents published between 2013 and 2020. Many collected documents makes the text mining an appropriate method for a thorough assessment of the literature (Cosic et al., 2015; Rikhardsson & Yigitbasioglu, 2018).

The map of the co-occurrence terms and extracted themes show that, although the research interest in data science is contemporary, the studied subjects cover a wide variety of knowledge areas. Our findings indicate that the knowledge structure mainly consists of research applied to data analysis on customers management that allow to forecast a possible churn and improve the satisfaction and loyalty of customers base, in general. Examining the publication data associated with the keywords in the co-word analysis network showed that the topics related to CRM systems have become more popular.

REFERENCES

Academic Acelerator. (2021). *Academic Accelerator: Accelerate your scientific research*. Retrieved from https://academic-accelerator.com

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. doi:10.1016/j.jss.2006.07.009

Buttle, F., & Maklan, S. (2019). Customer relationship management: concepts and technologies (4th ed.). Routledge, Taylor & Francis Group. doi:10.4324/9781351016551

Camacho, D., Panizo-Lledot, A., Bello-Orgaz, G., & Gonzalez-Pardo, C. E. (2020). The four dimensions of social network analysis: an overview of research methods, applications, and software tools. Information Fusion, 63, 88–120. doi:10.1016/j.inffus.2020.05.009

Cosic, R., Shanks, G., & Maynard, S. B. (2015). A business analytics capability framework. *AJIS. Australasian Journal of Information Systems*, *19*, S5–S19. doi:10.3127/ajis.v19i0.1150

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Coussement, K., Lessmann, S., & Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, *95*, 27–36. doi:10.1016/j.dss.2016.11.007

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. doi:10.1016/j.ejor.2018.02.009

De Caigny, A., Coussement, K., & De Bock, K. W. (2020). Leveraging fine-grained transaction data for customer life event predictions. *Decision Support Systems, 130*. . doi:10.1016/j.dss.2019.113232

de Castro Neto, H., Julia, R. M. S., Paiva, E. R. F., Carvalho, A. P. L. F., Junior, A. P. S., Peres, D. R. S., . . . de Assis, J. E. (2019). Improving the AHT in telecommunication companies by automatic modeling of call center service. In *EPIA Conference on Artificial Intelligence* (pp. 96–107). Vila Real, Portugal: Springer. 10.1007/978-3-030-30244-3_9

Delen, D. (2014). *Real-world data mining: Applied business analytics and decision making*. FT Press.

Delgosha, M. S., Hajiheydari, N., & Saadeatmantesh, H. (2020). Semantic structures of business analytics research: applying text mining methods. In *Information Research* (Vol. 25). University of Borås. Retrieved from http://www.informationr.net/ir/25-2/paper856.html

Feng, J., Zhang, Y. Q., & Zhang, H. (2017). Improving the co-word analysis method based on semantic distance. *Scientometrics*, *111*(3), 1521–1531. doi:10.100711192-017-2286-1

Iwashita, M. (2019). Transitional Method for Identifying Improvements in Video Distribution Services. *International Journal of Networked and Distributed Computing*, *7*(4), 141. doi:10.2991/ijndc.k.190911.001

Jussila, J., Huhtamäki, J., Kärkkäinen, H., Aho, T., Kortelainen, S., & Tebest, T. (2014). New era of business analytics - Making sense of business ecosystems. *MINDTREK 2014 - Proceedings of the 18th International Academic MindTrek Conference*, 266–268. 10.1145/2676467.2676517

Laudon, K. C., & Laudon, J. P. (2020). *Management information systems: managing the digital firm* (16th ed.). Pearson Education Limited.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal*, *9*, 181–212. doi:10.28945/479

Moher, D., Liberati, A., & Tetzlaff, J. (2015). Principais itens para relatar revisões sistemáticas e meta-análises: A recomendação PRISMA. *Epidemiologia e Serviços de Saúde: Revista do Sistema Unico de Saúde do Brasil*, *24*(2), 335–342. doi:10.5123/S1679-49742015000200017

Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, *29*, 37–58. doi:10.1016/j.accinf.2018.03.001

Rivera-Castro, R., Pilyugina, P., Pletnev, A., Maksimov, I., Wyz, W., & Burnaev, E. (2019). Topological data analysis of time series data for B2B Customer Relationship Management. https://arxiv.org/abs/1906.03956

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Salas-Rueda, R.-A., & Salas-Silis, J.-A. (2019). WeVideo: ¿Servicio en la nube útil para los estudiantes durante la construcción y presentación de los contenidos audiovisuales? *Vivat Academia*, *1*, 67–89. doi:10.15178/va.2019.149.67-89

Saur-Amaral, I., Ferreira, P., & Conde, R. (2013). Linking past and future research in tourism management through the lens of marketing and consumption: a systematic literature review Investigação passada e futura sobre gestão do turismo, marketing e consumo: uma revisão sistemática da literatura. *Conde / Tourism & Management Studies, 9*(1), 35–40.

Semrl, J., & Matei, A. (2017). Churn prediction model for effective gym customer retention. *Proceedings of 4th International Conference on Behavioral, Economic, and Socio-Cultural Computing, BESC 2017,* 1–3. 10.1109/BESC.2017.8256385

Sung, C., Zhang, B., Higgins, C. Y., & Choe, Y. (2016). Data-driven sales leads prediction for everything-as-a-service in the cloud. *Proceedings - 3rd IEEE International Conference on Data Science and Advanced Analytics, DSAA 2016*, (May), 557–563. 10.1109/DSAA.2016.83

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, *14*(3), 207–222. doi:10.1111/1467-8551.00375

Van Eck, N. J., & Waltman, L. (2011). *Text mining and visualisation using VOSviewer*. https://arxiv.org/pdf/1109.2058

Waller, M. A., & Fawcett, S. E. (2013). Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. *Journal of Business Logistics*, *34*(4), 249–252. doi:10.1111/jbl.12024

Wang, C. H., & Lien, C. Y. (2019). Combining design science with data analytics to forecast user intention to adopt customer relationship management systems. *Journal of Industrial and Production Engineering*, *36*(4), 193–204. doi:10.1080/21681015.2019.1645051

Wang, X., Nguyen, K., & Nguyen, B. P. (2020). Churn prediction using ensemble learning. In *ACM International Conference Proceeding Series* (pp. 56–60). 10.1145/3380688.3380710

Yue, B. (2020). Topological Data Analysis of Two Cases: Text Classification and Business Customer Relationship Management. *Journal of Physics: Conference Series*, *1550*(3), 032081. Advance online publication. doi:10.1088/1742-6596/1550/3/032081

KEY TERMS AND DEFINITIONS

Customer Acquisition: Business tactics aiming to attract prospect customers and to let them know about the enterprise' products and/or services and, the consequently, to identify uniquely the new customers when performing the purchases.

Customer Relationship Management (CRM): Corporate information management strategy inherent to customer relationship processes, supported by digital technologies, and aimed at increasing the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

efficiency and effectiveness of processes and respective activities, carried out with a view to acquiring new customers and retaining valuable customers' existing.

Customer Relationship Management System: Set of interdependent components—hardware, software, database resources, telecommunications networks, and people and procedures—that interact to collect data from customers and their interactions with the enterprise at the level of the various organizational functions - marketing, sales and support -, storage them, process them (transformation and analysis), and disseminate information to improve decision-making processes in order to acquire and retain valuable customers.

Customer Retention: Business tactics used by the enterprise to develop the relationship with new customers and improve their satisfaction level through interactions with the enterprise and, consequently, increase their value and avoid the future churn.

Data Science: Recent popular field related to business analytics research and its intelligent techniques applied to big data.

Systematic Literature Review: A systematic approach following a specific methodological process to review a set of academic literature, in order to analyses and discover new tendencies in a research topic.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 26
Data Science in Economics and Business:
Roots and Applications

Mara Madaleno
https://orcid.org/0000-0002-4905-2771
GOVCOPP, DEGEIT, Portugal & University of Aveiro, Portugal

João Lourenço Marques
https://orcid.org/0000-0003-0472-2767
GOVCOPP, DCSPT, Portugal & University of Aveiro, Portugal

Muhammad Tufail
https://orcid.org/0000-0002-6013-9157
National University of Sciences and Technology (NUST), Pakistan

ABSTRACT

Economics and business are a great background for data science provided econometricians and data scientists are sets with an intersection, although remaining unknown. In econometrics, data mining is somewhat a monstrous word, a field that traditionally seeks causal inference and results in interpretability. When we go deeper into what data science usually is, the boundaries between more traditional econometrics and even statistics and the hip and cool machine learning become shorter. In economics and business, we find examples and applications of simple and advanced data science techniques. This chapter intends to provide state-of-the-art data science applications in economics and business. The review and bibliometric analysis are limited to the research articles published through Elsevier Scopus. Results allowed the authors to conclude that despite the number of already existent research, a lot more remains to be explored joining both fields of knowledge, data since, and economics and business. This analysis allowed the authors to identify further possible avenues of research critically.

DOI: 10.4018/978-1-7998-6985-6.ch026

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION AND FRAMEWORK

Undoubtedly, mathematical modeling and statistics are central to quantitative economics, highlighted by the fact that some of the biggest data repositories are maintained by economic research organizations. Economists try to understand human behavior through Homo economics or the "economic man", to model human behavior. For that, they rely on economic theories and use some analytical tools and techniques, many of which are no more than standard statistical and mathematical models. We find in the literature economists walking around in all domains; from unemployment and inequality to the economics of climate change, to advertisement and revenue collection, from consumer and production behavior, up to profit, sales, and purchases, and finance data explorations (Provost and Fawcett, 2013; Brooks et al., 2019; Nosratabadi et al., 2020a, 2020b; Taddy, 2019; Basdas and Esen, 2020; Consoli et al., 2021).

We may think of some reasons why we should not dissociate economics, business, and data scientists. At least economists know exactly what the term machine learning means but with more econometric and statistical concepts. Machine learning ends up being a fancy word to describe statistical or predictive modeling that is used by programmers (Cao, 2017). Surprisingly, when we look for the most popular machine learning courses, the first two modules of their syllabus are still linear regression and logistic regression, basic concepts learned by economists in introductory econometrics. Econometrics is the application of statistical modeling in understanding complex social and environmental issues, being a big area of applied economics. We may even argue that in this case, economists have a deeper knowledge of linear regression than the average data scientists abroad. However, it is still hard for an economist to understand concepts such as neural networks, machine learning, and support vector machines, among many others, and still, heteroscedasticity is something away from the syllabus of machine learning, putting forward the pros and cons for economists or managers desiring to be involved in data science (Basdas and Esen, 2020; Consoli et al., 2021). Neural networks combine layers of logistic-like regressions to model non-linear relationships among variables, usually more complex, that simple regression analysis is not able to capture. Therefore, nothing as difficult as it may initially seem. Still, they build powerful algorithms, in what we believe to be complementary areas.

Economists have higher standards given their obsession in finding causal relationships, a goal unable to be fulfilled unless randomized controlled trials are pursued and sensitivity analysis to models basic assumptions scrutinized (we may think here of many emerging biases like attenuation, survivorship, selection bias, measurement error, reverse causality, truncation, censoring, among others). For machine learning, things are easier, given that models are not solved explicitly, and they do not need to rely on stricter assumptions (Donoho, 2017). Instead, models are estimated interactively with gradient descent and its derivatives, usually ignoring the theory behind the relationship we are trying to study (Taddy, 2019; Nosratabadi et al., 2020b). Even so, they use cross-validation and testing, and instead of t-statistics bootstrapping. In reality, this is summed up to the use of the right tools in the correct applications. Moreover, presentation and writing are important parts of data science provided non-technical audiences will be the interesting parts in research (managers, policymakers, marketers, copywriters, customers, clients). After all, the final result expected to be driven from the research is that we can demonstrate why our results matter and how stakeholders can use them and act over them. Economists, have generally a broader picture of the overall panorama and deepen these results, not just reading numbers as most data scientists, but presenting and clearly explaining results, with fluid writing. Whereas mathematicians and a computer scientist are not so comfortable when it comes to presenting and explaining the work

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

results clearly and to a broader audience (with a correct structure and logic behind the story economists need to tell).

Economists should make great data scientists but usually opt for more traditional fields like finance, academia, enterprises, institutions, governments, and so on, and in all these fields we have data science applications as will become evident in our results section. However, if economists have a great background and unique blend of technical, statistical, soft skills, and human skills, perhaps, data science positions would benefit from having econometricians doing the job, taking advantage of their awareness of possible shortcomings emerging from data mining and resilience in trying all that might work to answer to specific problems (Taddy, 2019; Nosratabadi et al., 2020b). The joint work of both would increase the skills of both and allow to resolve more specific issues quickly, efficiently, and effectively (Consoli et al., 2021). At least we see in the literature that recently the combinations of the different areas of knowledge (economics, management, data science) have allowed great analyzes recognized by pairs of problems that both sciences had hitherto tried to solve in a more individualized way.

In fact, in the areas of economics and business, we may find several examples and applications of simple and advanced data science techniques. For example, behavioral economists can quantify consumer response and attitude in response to marketing campaigns, price reduction, and addition of new features (Basdas and Esen, 2020). Behavioral finances also deal with consumer behaviors and attitudes but are related to financial markets, where consumers are called investors (Consoli et al., 2021). Financial markets have huge datasets with which to work, demanding simply more sophisticated econometric techniques to analyze this data (Brooks et al., 2019). In all activity sectors, we find applications of data science (Cao, 2017), namely, agricultural and energy economists help decide how to set the right insurance policy for farmers against natural disasters and to deal with energy efficiency, consumption, and implementation of the rights policies to pursue the low carbon footprint goals and energy transition goals (Nosratabadi et al., 2020a). Industrial economists determine the level of commodity production and pricing mechanism for given market demand, helping markets to define the equilibrium price to be applied and sports economists are already making a huge influence (Consoli et al., 2021). All these examples are explored through optimization models, theoretical models, and econometrics, turning evident the relationship between data science and the areas of economics and business, as also exposed above.

To quantify marginal changes in output for a small change in input, mathematical models like differential equations are used. Statistical models are typically used to develop and test theories on the relationship between economic agents (people, farms, businesses) and impacts of policies/incentives using statistical tools. The whole domain of "econometrics" is based upon the application of standard statistical models (Consoli et al., 2021). These are familiar to data scientists since we are talking about descriptive statistics and hypothesis testing, correlation and regression analysis, time series analysis and visualization, predictive analytics and forecasting, and, as well, panel data modeling (OLS – ordinary least squares, fixed and random-effects models, etc.). Provided most economic questions are empirical, standard scientific processes are applied to analyze it (developing a hypothesis, collecting data, testing the hypothesis, making conclusions) (Basdas and Esen, 2020).

This chapter intends to present a literature survey and bibliometric analysis of the applications of data science to business and economics, considering as analysis period the most recent twenty years initially, to draw attention to the growing body of application and critically analyzing avenues of future research. We do that by exploring methods and techniques, analytics toolbox, and highlighting specific case studies. Almost all economics is data-driven and quantitative. As such, a model of the real world is created (a small or large part of it as representative) to understand relationships between different com-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Data Science Glossary
Source: Bajracharya (2018).

ponents, their interactions, and the impact of external influences, like policies, incentives, and shocks. Thus, when this complex interconnected system is modeled, economists rely on a range of mathematical and statistical techniques.

From the analysis undergone, it was possible to infer that the amount of data available reached incalculable dimensions with the advancement of the information systems, so their treatment and analysis need to support companies' strategic decisions. Data science is the science of data (Cao, 2017). Big data is only one of its dimensions, referring to data that is too large and/or complex to be effectively and/or efficiently handled by traditional data-related theories, technologies, and tools. Big data is changing the traditional business models because companies created in the digital era are aware of the potential of big data, and understand their data, how to use it, and the opportunities for gaining competitive advantage (Donoho, 2017). Big data is helping managers to measure and manage and as a result, they know more about their business, and this knowledge is translated into an improved decision-making process and increased performance (McAfee et al., 2012).

Economics and business are therefore much more than programmers think, as it is data science for economists or managers. Figure 1 presents a word cloud of data science key terms providing a clear picture of all that is involved.

Literature includes an appropriate number of state of the art review papers and comparative analysis of many data science methods (e.g., Brooks et al., 2019; Nosratabadi et al., 2020a; Nosratabadi et al., 2020b; Basdas et al., 2021). Considering the fields of economics and business jointly, there is still a lack

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of research in the area and the current research aims to fill this gap by providing a recent bibliometric and brief bibliographic analysis of methods within data science used and applied for economics and business. The chapter further contributes to identifying future trends in the advancement of learning algorithms for the economics and business areas.

The rest of the chapter develops as follows. Section 2 presents the data and the methodology followed to present the bibliometric and bibliographic analysis. Section 3 presents the results and discusses them, while section 4 evidences some of the most relevant conclusions.

DATA AND METHODOLOGY

Given that our goal is to provide the state of the art of data science in economics and business, the review and bibliometric analysis are limited to the original peer-review research articles published through Elsevier Scopus, which is one of the most comprehensive journals databases in the world. Besides, only articles written in English are considered for the review in this study.

The systematic literature review and bibliometric approach were based on the Prisma method including four steps: 1) identification (documents are identified through an initial search in the mentioned databases), 2) screening (first, duplicate articles are eliminated; second, the relevance of the articles is examined), 3) eligibility (the full text of articles was read by the authors, filtered and only some of them were considered eligible), 4) inclusion (the creation of the final database of the study, which is used for qualitative and quantitative analysis).

It was possible to observe that application domains include stock market, marketing, E-commerce, corporate banking, and cryptocurrency, but many more. Finally, we limited the search to the years 2018-2021, thus three complete years of data, as many articles had been already accepted to be published in 2021, or were in the publishing process. Removing other documents and letting solely articles and book chapters, we ended up with 377. Two different approaches were followed for the analysis of the collected data (articles): i) the VOSviewer, a bibliographic analysis software (https://www.vosviewer.com/), is used to visualize and analyze some bibliometric features of the initial data (N=377), based on the principle of co-occurrence and coupling of literature; and for ii) the most relevant publications (N=13), those identified as having the highest number of citations (see Table 3), a content analysis was applied to identify how the different topics studied in Economics and Business are linked with data science approaches. For this purpose, the final documents were stored as text files in QDA Miner and using the WORDSTAT software (developed by Provalis Research: https://www.provalisresearch.com/) a cluster analysis is presented for the key phrases of these texts. Once these groups of key expressions were found, the "key-word-in-context" (KWIK) tool was used to better understand the meaning of these contents in their contexts. The outputs of this analysis are presented and analyzed in the next section.

As for the search strategy of articles in the Scopus database, our initial research methodology was very general. While looking for **Data Science in Economics and Business** in title abstract and keywords Scopus returned 3,168 results. Restricting research only for the years 2011-2021 yielded 2,557 document results available. Limiting the search for the subject areas Economics, Econometrics and Finance, and Business, Management, and Accounting we ended up with 1,065 document results. From these, 894 are articles, 47 conference papers, 39 book chapters, 30 books, and 30 reviews (the rest relies on erratum, conference review, editorial, and letter and retracted). We then restricted the search solely to articles and book chapters, yielding 933 documents (both in the final publication stage as well as articles in press).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

To restrict it further we limited the publishing period to the last 5 years (not considering 2021 since it is not full), 2016-2021, and limited our search to solely published articles in Journals. Even so, the number of retrieved articles was high dimension. Therefore, we advanced to other search criteria inside the Data Science umbrella, considering only those articles which in their title, abstract, or keywords, included the words *big data*, *data mining, data handling, artificial intelligence*, *digital storage*, *machine learning*, *information management*, *cloud computing*, *big data analytics*, *big data applications* and *data integration,* of course, limited to the subject areas of Economics, Econometrics, and Finance, and Business, Management, and Accounting.

The search limited to the years, 2016-2021, peer-reviewed articles and published or accepted to be published in scientific journals, under the general umbrella Data Science in Economics and Business, allowed us to end up with 2,235 document results. Curiously between 2018 and 2019, the number of articles published almost doubled. Published articles by year under the search limited to years 2016-2021 and for the retrieved 2,235 documents yielded 199 in 2016, 347 in 2017, 369 in 2018, 611 in 2019, 618 in 2020, and in 2021 we already have 91 documents published corresponding to the exposed criteria.

Further filters were applied to this search, namely in terms of years, provided the initial search criteria were still retrieving many articles. Therefore, we limited the articles between 2018-2021. These yielded for Business, Management, and Accounting 369 documents, and Economics, Econometrics and Finance 167, and for Social Sciences Scopus retrieves 109 articles, in a total of 430 documents. Provided it was still a lot of articles to analyze, we decided to restrict the analysis further, but surprisingly it is vast and huge the existent applications of data science in business and economics. Using the previous search, we limited articles further to the following subjects: *big data*, *economics*, *business,* published between 2018-2021, *data mining*, *data handling*, *artificial intelligence*, *digital storage*, *machine learning*, *information management*, *cloud computing*, *big data analytics*, *big data applications*, *data integration*, and excluding all subject areas besides Business, Management and Accounting, and Economics, Econometrics, and Finance. Removing other documents and letting solely articles and book chapters we ended up with 377 documents. Restricting only to the most recent years of 2018-2021 was only an attempt to restrict the total number of articles retrieved. Figure 2 summarizes the systematic process followed to select the literature.

The next section will present the results of our bibliometric and bibliographic analysis considering these 377 results. From these 363 refer to a journal published articles and only 14 to book chapters. The inclusion of book chapters in our research is because we are exactly elaborating this analysis for a book chapter inclusion and as such we taught it would be relevant to include them as well.

RESULTS AND DISCUSSION

In this section, a bibliometric analysis for the 377 papers, extracted from the Scopus database limited to the years 2018-2021, is presented. This analysis was conducted using the VOSviewer (version 1.6.13) software to retrieve and visualize the most relevant information about the network structure in these topics, such as co-occurrences of the author keywords, words in the titles and abstracts of the publications, and the bibliographic coupling of the authors, countries, institutions, and journals. Finally, a co-word analysis was carried out, using the QDA miner (version 6) and WORDSTAT (version 7.1) tools, to highlighting the crucial dimensions of the phrases network structures of the most relevant full papers (N=13). First, to this presentation, we have concentrated the discussion in the analysis of the method-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Systematic process to select literature

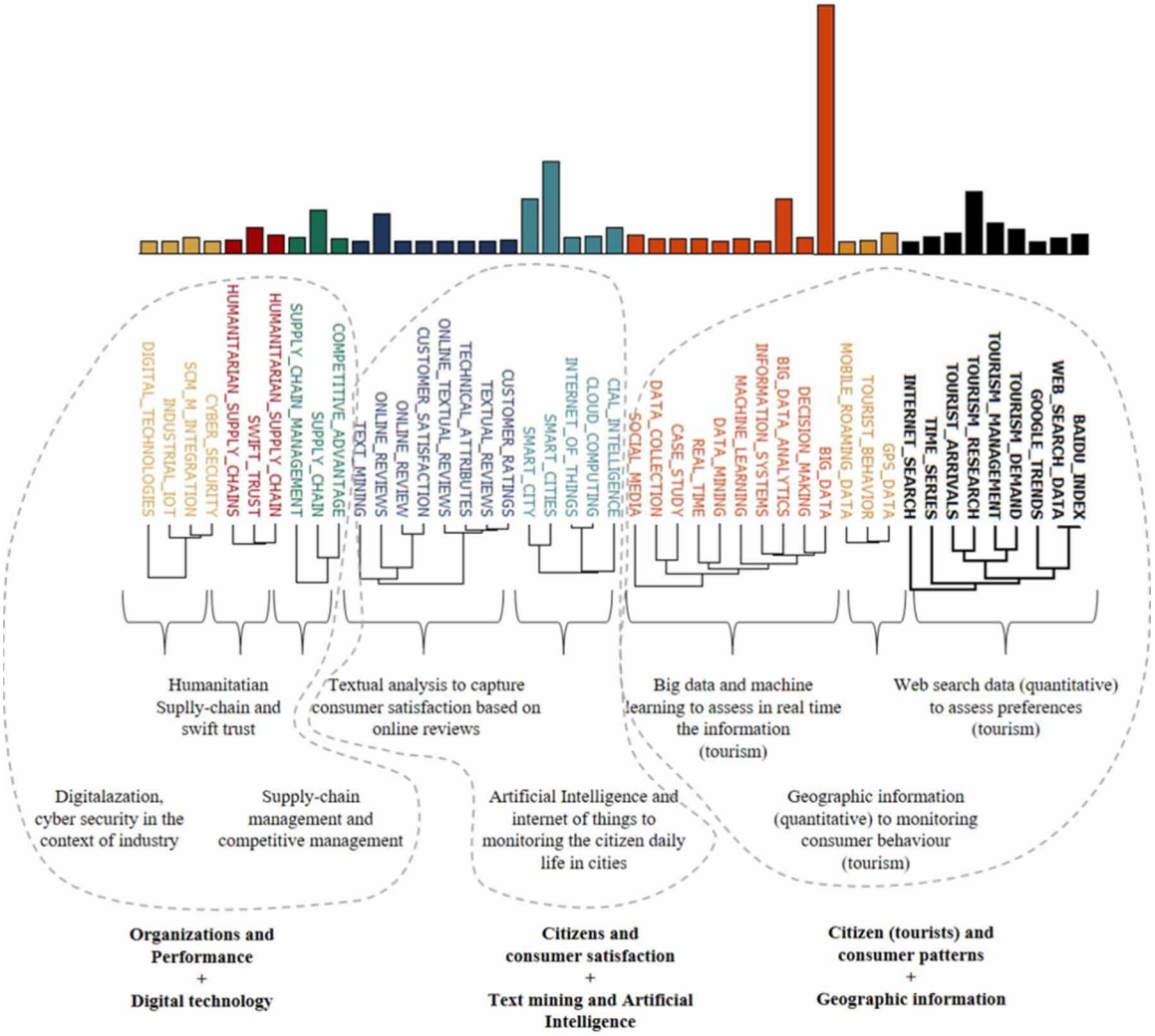
ologies followed, in summary, by the 377 articles analyzed (in a kind of bibliometric analysis for the employed methodologies and methods).

The first step of this analysis is to show the evolution of the number of publications in the last five years (2018-2021). There has been an increase of published articles under the subject areas of economics and business and any data science subject. The year 2021, still at the beginning (middle of February), has already a significant number of publications, meaning that a further increase is expected by the end of this year (see Figure 3).

To support the bibliographic analysis indicated previously as part of this chapter development, we resorted to the entire database of articles to explain which methods have been mostly applied from Data Science techniques to Economics and Business/Management areas analysis. We can take from the different collected articles in the analysis period that deep learning and machine learning methods are useful contributions in model advancements in prediction, planning, and uncertainty analysis. Therefore, with a wide application in economics, econometrics, management, business, and finance. Among deep and machine learning methods applied to address the different aspects in these area subjects, we find in the literature applications of artificial neural networks, support vector machines, decision trees, ensembles, Bayesians, hybrids, and neuro-fuzzy.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Published articles by year under the search limited to years 2018-2021 and for the retrieved 377 documents



These have been applied in the articles analyzed, being easy to understand that the emerging technologies and novel concepts which arose in several fields of education and day life aspects. These have been promising to encourage a brighter future in dealing with these aspects and to provide systems to decision support. It is pointed in the literature that artificial intelligence applications allow us to use the internet of things, machine learning, deep learning, and big data to support the evolution and technological advancement in the economics and business fields. These methods have widely contributed to various applications domains with promising results as evidenced by the previous bibliometric analysis presented (e.g., smart cities, mobility, management, monitoring, city planning, resources allocation, energy demand, sustainability, corporate bankruptcy, consumption prediction, consumption reviews, food supply, cryptocurrency, production prediction, financial forecasts, air pollution monitoring and prediction, exchange rates, tourism demand, and supply, etc.). As for the future, it is further expected that the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

trends will converge towards the evolution of sophisticated hybrid deep learning models in economics and business subjects and fields.

Comprehensive state-of-the-art research is provided by Nosratabadi et al. (2020b), exposing the recent developments in emerging economic applications. Their analysis is done regarding novel data science methods, particularly, deep learning, machine learning, hybrid deep learning, hybrid machine learning, and ensemble models. They mention that broader application domains include economics research from stock market behavior, marketing, e-commerce, corporate banking, and cryptocurrency. In the present work we confirm these findings but as well join these tourism consumer reviews and forecasts, financial management, smart cities planning and management, consumption and production management, and consumer reviews and forecasts, despite many others.

From the bibliographic analysis performed it was as well possible to take many insights. Machine-learning algorithms provide the ability to learn from data and deliver in-depth insight into problems (Nosratabadi et al., 2020b). Notable classic machine learning methods (data science) applied in economics are support vector regression for anomaly detection, naïve Bayesian and C4.5 decision tree classifiers for credit card fraud detection, improved BP neural networks for product export and volume prediction, multilayer feed-forward neural network for stock price predictions, multilayer feed-forward neural network for stock price prediction, artificial neural networks for mobile social commerce, adaptive neuro-fuzzy inference systems for e-banking failure, decision trees for customer behavior, deep neural networks for recommender systems, naïve Bayes and linear support vector machine for sentiment analysis, deep learning for citizens and network analysis and tourism and forecasting, among many others (Brooks et al., 2019).

Moreover, deep learning as an emerging field of machine learning was applied mostly to aspects of today's society, self-driving cars, image recognition, animal behavior, renewable energy forecasting, hazard prediction, health informatics, bioinformatics, and so on. Deep learning models are usually evaluated in economics by support vector machine, K-nearest neighbors, and generalized regression neural networks. Furthermore, hybrid machine learning models are formed through the combination of two predictive machine learning algorithms or by a machine learning algorithm combined with optimization methods. Therefore, it is used to increase the accuracy of the other models. Finally, ensemble machine learning algorithms combine and apply multiple learning and training algorithms to enhance training and learning algorithms allowing us to learn from the data. Hybrid methods have been used to explore financial time series, forecasting, and rate prediction (Nosratabadi et al., 2020a, 2020b; Basdas et al., 2021).

Long short-term memory methods are notable machine learning and deep learning methods applied in stock markets (Basdas et al., 2021). These were used to explore portfolio management, algorithmic trading, socially responsible investment portfolios, stock price prediction, exchange-traded-fund options, price prediction, automated stock trading, sentiment analysis, and market index and stock trend predictions. Notable machine learning and deep learning methods in marketing and tourism include as goals to explore customer behavior, promotional activities, consumer reviews and opinions, sentiment analysis, and opinions forecast. As to e-commerce, data science techniques have been mostly used to explore investment quality, evaluation, order arrival prediction, allocation issues, credit risk evaluation, and product recommendation. In cryptocurrencies exploration, we mostly observe that the research objective has been focused on cryptocurrencies price prediction, while corporate bankruptcy applications are mostly restricted to prediction.

Still, a lot more in Economics and Business should be followed, explored, and done. The 377 articles were those retrieved from the Scopus database after the filters applications. Taking advantage of Big Data, deep learning, and machine learning methods, many other issues and thematics could be explored

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in Economics and Management. As for now, we could only say that this is an emerging field of research and we should bear in mind the need to keep on the right track.

Bibliometric Analysis of the Keyword Provided by Authors

Figure 4 shows the network visualization of the co-occurrence of the keywords. The high number of expected nodes imposes a minimum threshold of words. In this case, only were considered the keywords provided by authors of the paper that had a frequency more than 2 times, meaning that, of the 1542 words, 54 met the threshold. The concept '*big data*', '*big data analytics*' and '*artificial intelligence*' were the most frequent keywords, with 292, 57, and 42 occurrences, respectively.

Each keyword is represented by a circle and the size of a node determines the frequency of its occurrence. The larger the circle, the more often the keyword appears. Analyzing the graph is also possible to assess the connections between those terms, i.e., how they co-occurred as keywords. The distance between the nodes indicates the proximity of the two keywords. The keywords with the highest total link strength were selected and for these three most frequent keywords the number of co-existence was 340, 82, and 70, respectively - a minimum of the strength of 2 is considered in Figure 4 for more clarity.

Six different clusters are presented, resulted from the level of the interconnectivity of the keywords. These major groups are identified and classified according to the object of analysis (marketing and e-commerce, citizen and tourism, industry and firms) and the kind of topics typically associated with the field of data science (big data and artificial intelligence, network analysis, and forecasting, digitalization and data security).

Analyzing Figure 4, there is a clear association between i) network and text analysis on topics related to human behavior (social media, tourism, and smart cities); ii) artificial intelligence (and more specifically machine learning mechanism applied to big data to analyze marketing issues and forecast consumer behavior; and finally, iii) digitization and process automation in the business and business context. In this last group, intellectual capital, information and cybersecurity, and data privacy are highlighted aspects.

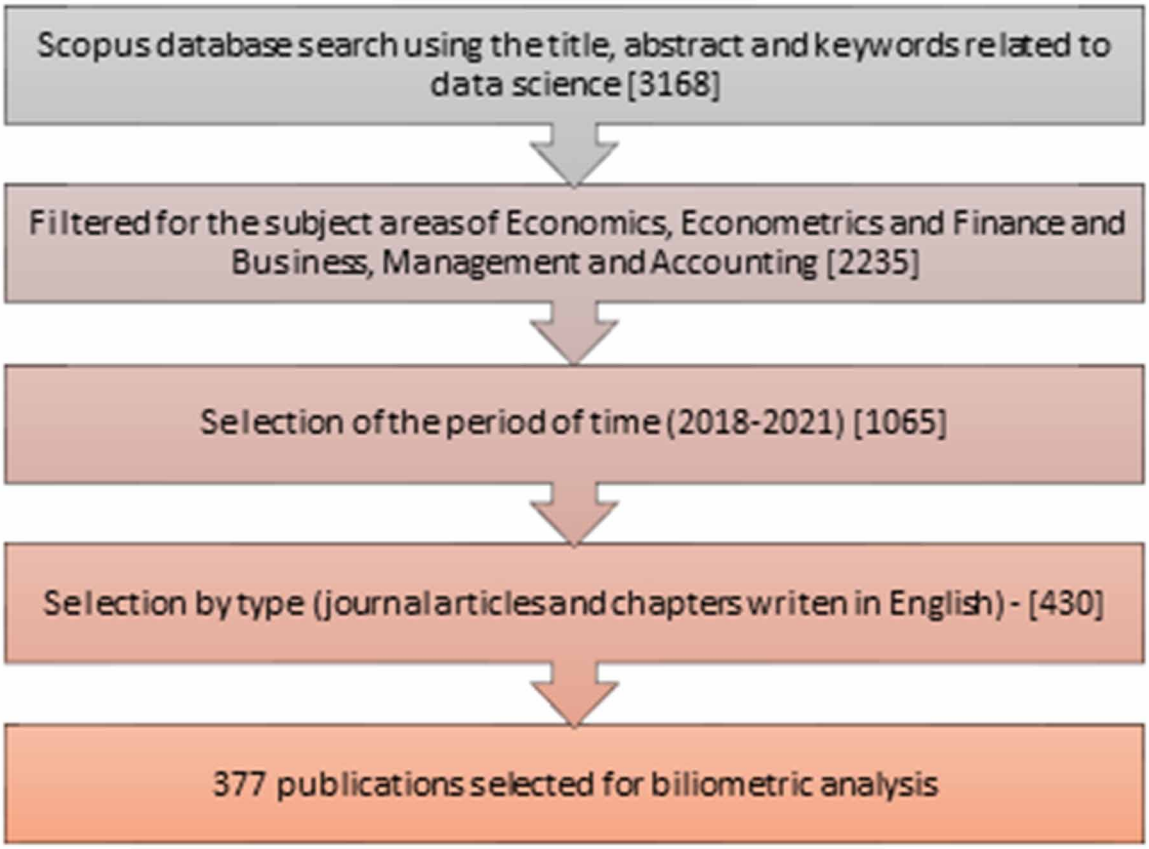
Bibliometric Analysis of the Words in the Title

Figure 5 presents a similar exercise to the one presented above, but this time applied to the titles of publications. Thus, the idea was to extract the concepts that occurred more than 2 times and group them according to their proximity - of the 1133 word 146 met the threshold. These results allow a more detailed analysis of the relationship between the themes and the responses that are intended to be obtained.

A total of 18 clusters were identified, however, we highlight 5 groups: i) the one that relates the services and the assessment of its consumers (applied with great frequency in the tourism area); ii) in another group, the processes of digitalization in firms and organizations stand out; iii) a third group, the relationship between performance assessment and monitoring in business contexts is evident; iv) it is also noteworthy, as already highlighted in the previous analysis, the application of a machine learning processes to assess consumption practices; and finally, v) the last group, aspects of perception assessment arise with the use of large volumes of data.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. The network visualization of the keyword provided by authors

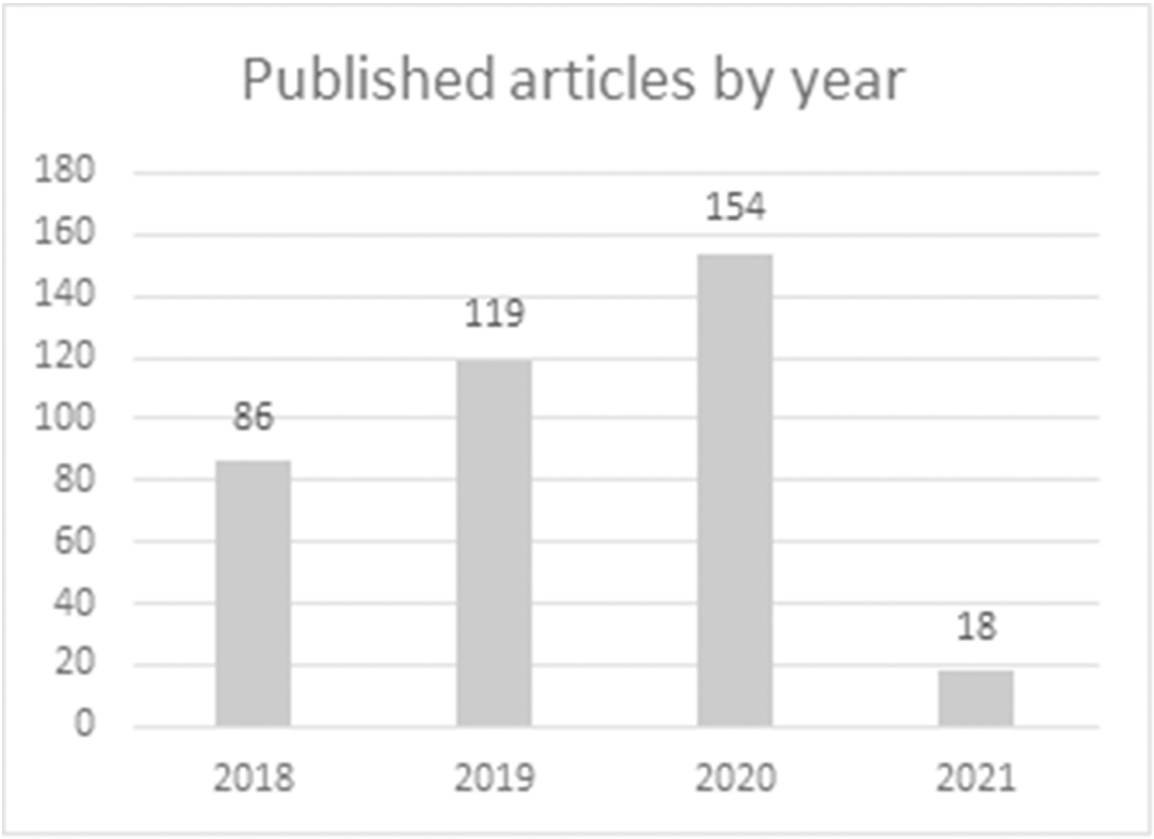


Bibliometric Analysis of the Words in the Abstracts

The analysis presented in Figure 6 highlights the most pertinent concepts included in the abstracts. These results stand out from the previous two analyses, due to the greater detail in the input data, using for this purpose the data of the 377 abstracts. The criteria of selection were the words or expressions within a sentence that occurred more than 15 times - of the 8402 words, 109 met the threshold (the Figure only shows the connections with a minimum strength of 10). The results are in line with the main conclusions of the previous analyzes (keywords and titles). Despite the greater complexity of the input data, three distinct clusters are presented: i) the first group is related to the adoption of performance evaluation mechanisms in the organizational management processes; ii) then, it is possible to highlight the analysis of consumption patterns taking advantage of big data generated, for example, from the online services and e-commerce; finally, iii) the third group aggregates content in the domain of economy and business in which the use of artificial intelligence to assess and estimate trends stand out.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The words visualization in the title



Bibliometric Analysis of the Authors

In total, 1023 authors were involved in the 377 articles analyzed. In Table 1 the most relevant author's names are revealed as to the number of articles published by each. Those with just one article published from the 377 have been excluded from the presentation. Only Dumay, J. has four articles published from the 377 from which resulted in our search. John C. Dumay is affiliated with the Macquarie University, Department of Accounting and Corporate Governance, thus a researcher working in the management field of education and not a data scientist (La Torre et al., 2018a; 2018b; 2019; Ndou et al., 2018). His co-authors only had three, two, or even one article published from these 377 during 2018-2021.

Bibliometric Analysis of the Institutions

Regarding affiliations of the authors publishing in the area of data science in economics and business, we observe from Figure 7 that eight are from the Hong Kong Polytechnic University, seven from the Macquarie University, and an equal number from the Università del Salento. Università Degli Studi di Torino has six affiliated authors among the 377 published articles in 2018-2021. It is important to note that the Macquire Business School is a part of The Macquire University. As such, in total there are 11 papers from this University (see Figure 7).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The words visualization in the abstracts

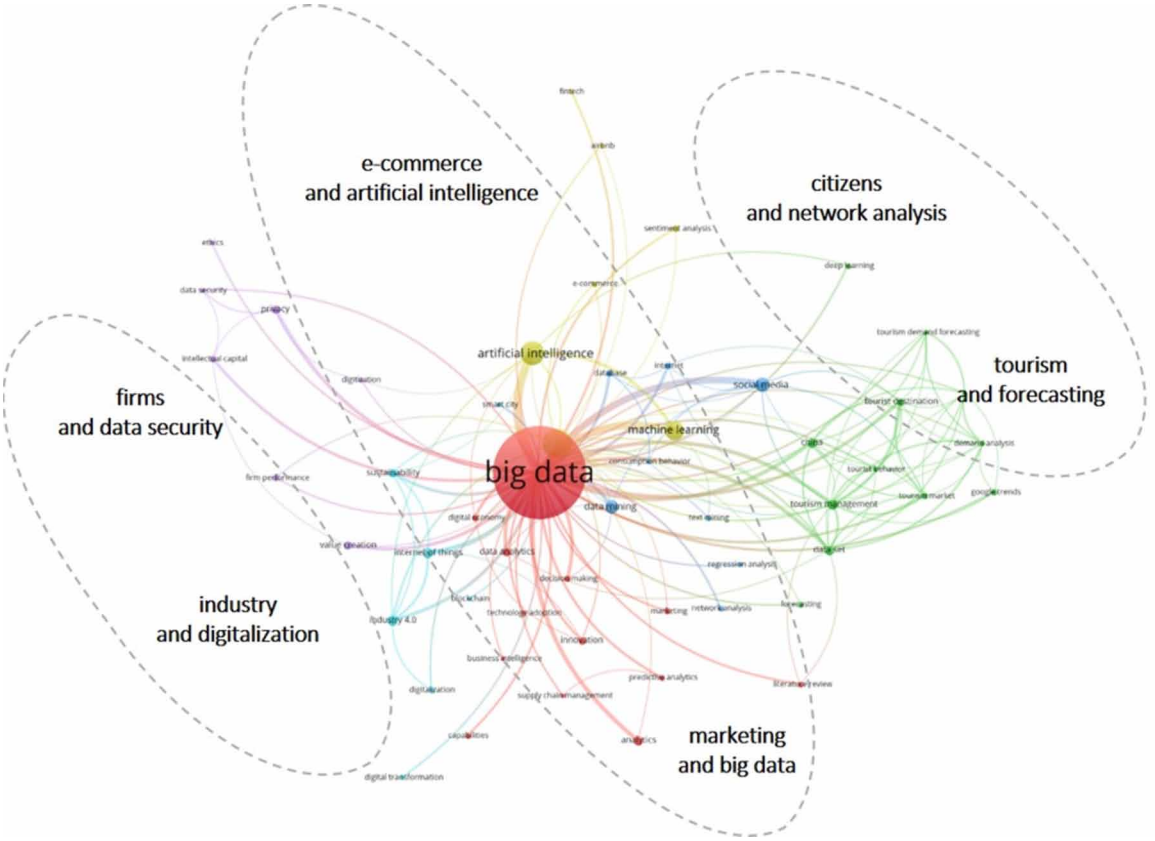
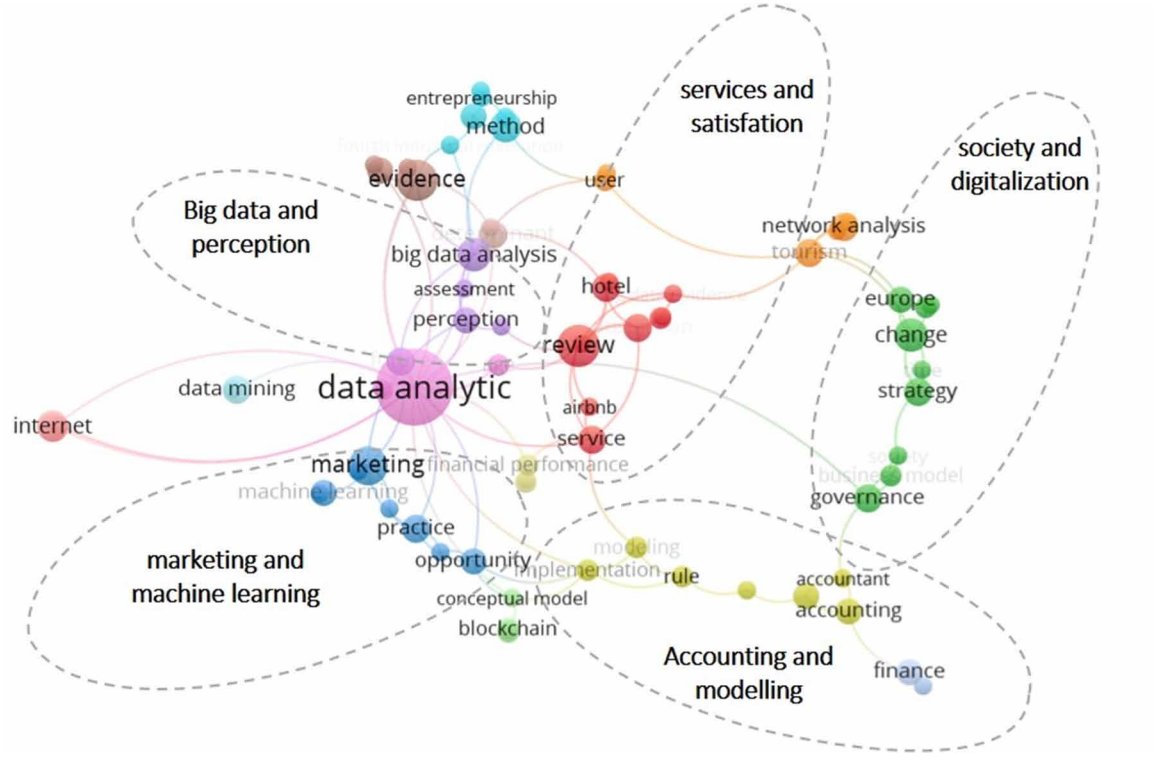


Table 1. Number of publications by author

Dumay, J.	4	Buhalis, D.	2	Gunter, U.	2	Ndou, V.	2
Akter, S.	3	Calic, G.	2	Hofmann, E.	2	Obschonka, M.	2
Kim, K.J.	3	Caputo, F.	2	Kaivo-oja, J.	2	Odendaal, E.	2
La Torre, M.	3	Chierici, R.	2	Kim, C.	2	Palmatier, R.W.	2
Law, R.	3	Choi, C.	2	Kim, K.H.	2	Park, Y.E.	2
Lim, C.	3	De Santis, F.	2	Kim, M.J.	2	Pellicelli, A.C.	2
Mandal, S.	3	Del Giudice, M.	2	Kirilenko, A.P.	2	Polimeno, G.	2
Passiante, G.	3	Del Vecchio, P.	2	Li, L.	2	Rea, M.A.	2
Sivarajah, U.	3	El Refae, G.A.	2	Maglio, P.	2	Rialti, R.	2
Tang, L.	3	Elia, G.	2	Maglio, P.P.	2	Roth, S.	2
Wang, S.	3	Fosso Wamba, S.	2	Mahroof, K.	2	Solazzo, G.	2
Agarwal, S.	2	Fosso Wamba, S.	2	Mariani, M.M.	2	Valentinov, V.	2
Anshari, M.	2	Gepp, A.	2	Marine-Roig, E.	2	Xiu, D.	2
Botes, V.L.	2	Ghasemaghaei, M.	2	Mikalef, P.	2	Zaki, M.	2
Brinch, M.	2	Gunasekaran, A.	2	Moro, S.	2	Önder, I.	2

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Affiliation of authors' published articles



Other strongly recognized worldwide universities emerge with four authors affiliated in each, where many others were removed from the graph with three, two, or even one author affiliated.

Bibliographic Coupling of the Countries

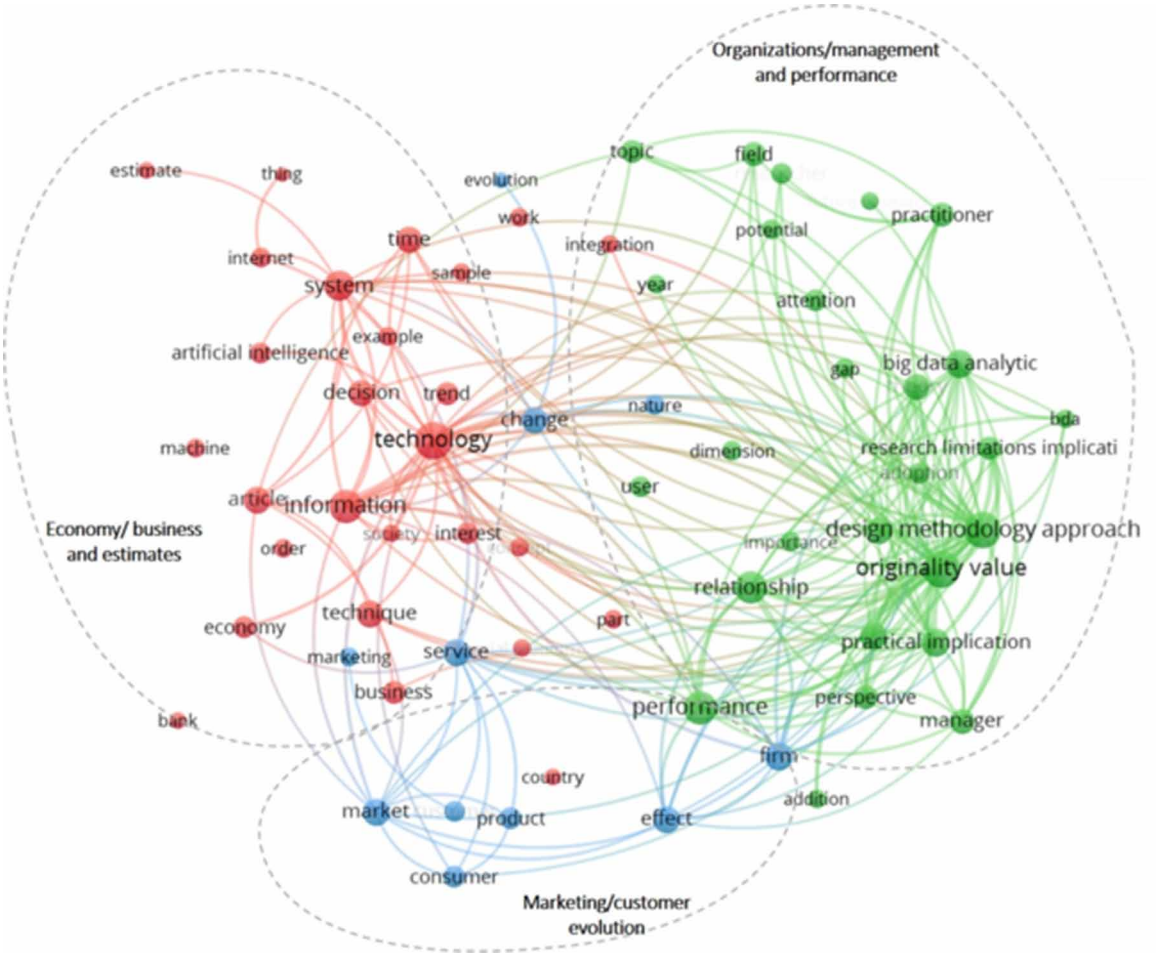
Figure 8 represents an analysis of articles published by countries or territory in the period between 2018 and 2021. Figure 8 (a) shows countries of sourced articles (for the 377 retrieved documents) up to several six in terms of affiliations. The United States appears at the forefront within the field, followed by the United Kingdom, Italy, China, and Australia. From Germany, we have 29 authors, whereas Portugal is in the sixteenth position, as 9, the same number as Norway.

Figure 8 (b) shows a biographic coupling of the countries, representing the link strength calculated based on the number of publications. A minimum of two publications by country was considered, meaning that 46 of 69 met the threshold. Different clusters that were more related (more frequently linked) are represented by colors. It means that the publications that originated from the countries in the same cluster have more co-authorship, repeatedly. The level of cooperation between countries, standing out 9 different groups: United States, United Kingdom, Italy, Germany, France, Finland, Russian Federation, United Arab Emirates, and Singapore are the countries that constitute the central nucleus of these clusters.

From here we could even establish a ranking among the world's top-ranking universities and affiliations of authors of the 377 documents retrieved from the Scopus search during this period. Still, research emerging within the field of data science for economics and management is associated with the most

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Country or territory of affiliations and network visualization
Note: (a) Country or territory of affiliations of authors' published articles. (b) The network visualization of the country cooperation during 2018 and 2021.



relevant countries in terms of development in research and science, and, of course, to the university reputation. Education within both fields is lacking in poorer countries opening room for deeper analysis and exploration, and an emerging opportunity for these countries' education concerning a joint analysis of data science techniques and economics and business. However, we should be aware of the hardest difficulties faced by developing countries in terms of education and this passage must be done carefully and gradually, provided more urgent needs in terms of economic development are still needed, just not to mention the hardness it is to get appropriate data to work with.

Bibliographic Analysis of the Journals

Exploring source title, or else journals names where these 377 articles have been published recently, provides a clearer picture of the subject areas which have been analyzed deeply in recent years (2018-2021). Table 2 presents the titles of the sources where articles have been published. With 19 articles

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

published, we have the Journal of Business Research published by *Elsevier Inc* under the subject area of *Business*, *Management and Accounting*, namely *Marketing*, being classified as Q1 journal. Also in the marketing category, we have the Industrial Marketing Management Journal with 17 of these 377 articles published, followed by Tourism Management with 13 articles, Economics, Management and Financial Markets Journal with 12 articles published, and the Business Process Management Journal with 10. Many more Journals are included in the search but excluded from the presentation in Table 2 since these had only 2 articles published in each (96 more journal titles and source titles, to be more specific).

Going deeper into the analysis of the words that are most representative within the total list of journal publications, Figure 11 provides the words list that most emerge from here. Just to mention a few, appearing 92 times we have the word Journal (which is expected considering the names of the journals, and as such ignored in the analysis), 39 times Management, 28 International, 25 Business, 23 Economics, 19 Marketing, 18 Research (as well ignored), 14 Review (also should be ignored), 11 Economic, 10 Tourism, 9 Accounting, 7 European, 6 Finance, 6 Financial, 5 Letters (ignored), 4 Hospitality, 4 Innovation, 4 Studies (ignored), 3 Auditing, 3 Econometrics, 3 Global, 3 Governance, and 3 Technology.

Therefore, it is more common to verify data science applications related to management or business issues (39 and 25, respectively), than to economics (23 and 11), although there are numerous tourism applications, surpassed by the number of publications within the finance subject (12), followed by accounting applications. The rest of the words just appearing once have been excluded from this theoretical analysis, but they are presented as well in Figure 9. Management, finance, and economics get out highlighted from this words cloud and most authors publishing within these journals are from the areas of economics and business and not specialized data scientists. However, there is still room for more data science applications, especially considering the big data available, and for example, in finance, we may find several opportunities for data analysis using as resources high-frequency data (for example, betting's, stock market prices, etc.). It is worth mentioning that from the subject fields the 377 documents search yielded for Business, Management, and Accounting 321 articles, for Economics, Econometrics and Finance 155, and under the general umbrella Social Sciences 103. This calls attention to the fact that these are interrelated topics and one Journal has more than one subject category.

From the 377 articles collected for this analysis, we went to explore, as well, the number of citations and the number of citations for each excluding self-citations. Results are presented in Tables 3, 4, and Figure 10. The h-index of citations was 28 meaning that of the 377 documents considered for the h-index, 28 have been cited at least 28 times between 2018 and 2021. Table 3 only presents those articles cited more than 50 times in total.

For the 377 articles, the total number of citations reported is 3,144 during the period of analysis considered, mostly in 2020 (1,916 citations), being the most cited article that of Wirtz, Patterson, Kunz, Gruber, Lu, Paluch, and Martins (2018), "Brave new world: Service robots in the frontline" published in the Journal of Service Management. It is also observed from this Table 3 that those articles which are mostly cited refer to the subject of tourism followed by applications to financial markets. This is due to the very nature of data regarding both fields. In tourism related to customer satisfaction opinions, online textual reviews, forecasting tourism arrivals, smart hospitality, among others. Smart cities is a well-explored subject and cited in the context of data science, perfectly justified by the huge amount of data available to be explored. In terms of financial applications although the financial performance application reveals to be the article with most citations, loan prospecting, sustainability and performance, production system and big data in industry 4.0, data abundance and asset price informativeness, financial statement fraud and data mining, high-frequency trading data, banks challenges in industry 4.0, digital

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

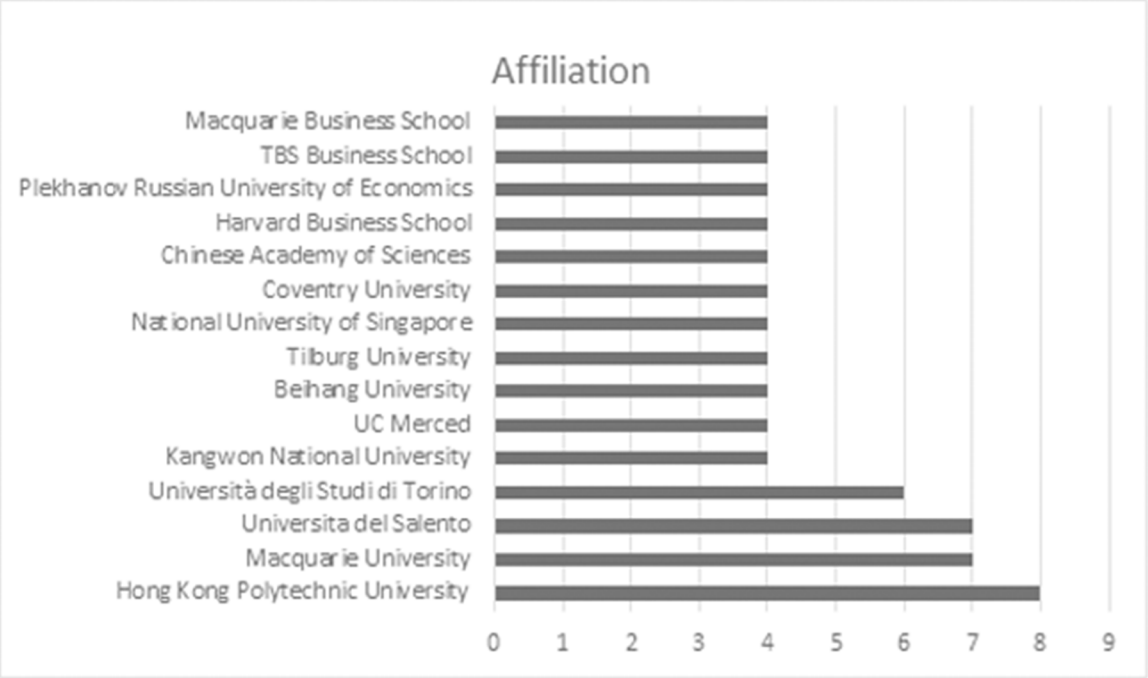
Table 2. Number of publications by source title or Journal

Journal Of Business Research	19	Asia Pacific Journal Of Tourism Research	3	International Journal Of Accounting And Information Management	2
Industrial Marketing Management	17	Business Horizons	3	International Journal Of Educational Management	2
Tourism Management	13	European Management Journal	3	Journal Of Accounting And Organizational Change	2
Economics Management And Financial Markets	12	International Journal Of Physical Distribution And Logistics Management	3	Journal Of Destination Marketing And Management	2
Business Process Management Journal	10	International Journal Of Research In Marketing	3	Journal Of Financial Economics	2
Meditari Accountancy Research	8	Journal Of Cultural Economics	3	Journal Of International Accounting Auditing And Taxation	2
International Journal Of Forecasting	7	Journal Of Knowledge Management	3	Journal Of Management And Governance	2
International Journal Of Logistics Management	7	Journal Of Service Management	3	Journal Of Monetary Economics	2
Journal Of Asian Finance Economics And Business	7	Journal Of Services Marketing	3	Journal Of Open Innovation Technology Market And Complexity	2
Journal Of Econometrics	7	Wseas Transactions On Business And Economics	3	Journal Of Sustainable Tourism	2
International Journal Of Contemporary Hospitality Management	6	Academy Of Accounting And Financial Studies Journal	2	Journal Of Urban Economics	2
Cities	5	Accounting Research Journal	2	Marketing Letters	2
International Journal Of Economics And Business Administration	5	Bottom Line	2	Research In The Sociology Of Organizations	2
International Journal Of Hospitality Management	5	Competition And Change	2	Research In Transportation Economics	2
Quality Access To Success	5	Current Issues In Tourism	2	Review Of Industrial Organization	2
Managerial Auditing Journal	4	Economics Of Innovation And New Technology	2	Risks	2
Marketing Science	4	European Competition Journal	2	Singapore Economic Review	2
Small Business Economics	4	Global Business And Economics Review	2	Technological And Economic Development Of Economy	2
Technology In Society	4	Information Resources Management Journal	2	Tourism Analysis	2
Tourism Economics	4	International Finance Review	2	Tourism Management Perspectives	2

systems and financial management challenges, credit scoring, commodity price risk, hidden sentiment in financial big data, financial markets efficiency, stock prices trends, forecasting financial cycles, artificial intelligence, and financial markets, are just a few of the several applications which have been presented in the form of articles published up to this moment.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Word cloud of Journals (articles published within), under the search, limited to years 2018-2021 and for the retrieved 377 documents



Given that the number of self-citations should be removed from an analysis of this type, Table 4 presents the number of citations reported during 2018-2021 for the collected 377 references, this time excluding self-citations. We followed the same criteria and considered only those with a number above 50.

As compared to Table 3 results, some articles are now left aside for not reaching the 50 citations, and some moved position. Curiously, the most cited article remains to be the most cited article without considering self-citations, even if with 10 fewer citations. Now in total for the period considered, the real total number of citations is just 2,796, less 348 citations which are self-citations. This total number considers all the 377 article citations overall, by year, and for the total period of 2018-2021. Finally, we have gathered all the references list of our 377 published articles and in Figure 11 we present a word cloud for the terms which emerge from the published articles, leaving room for a deeper analysis of the room that is still left for research under these thematic.

Big data analysis, business, tourism, marketing, performance, financial analysis, industrial, forecasting supply, digital, mining, smart systems, intelligence, learning, and models, are still the words that are commonly explored in data science applications to economics and business subjects.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Number of citations during 2018-2021 for the collected 377 references

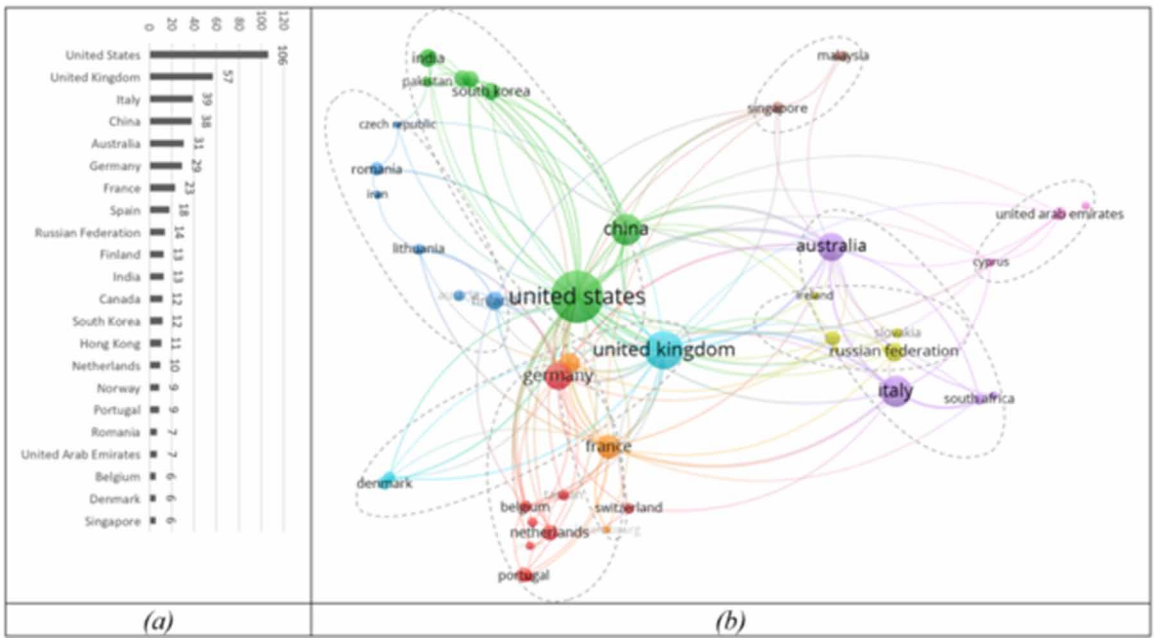
Publication info			2018	2019	2020	2021	Total
Year	Authors	Journal Title	105	740	1916	380	3144
2018	Wirtz et al.	Journal of Service Management	5	42	131	20	198
2018	Li et al.	Tourism Management	4	43	96	23	166
2018	Lim et al.	Cities	3	38	46	9	96
2019	Allam and Dhunny	Cities	0	14	62	11	87
2019	Zhao et al.	International Journal of Hospitality Management	0	14	52	12	78
2018	Amado et al.	European Research on Management and Business Economics	10	28	34	3	76
2019	Ardito et al.	Business Process Management Journal	0	14	48	10	72
2018	Buhalis and Leung	International Journal of Hospitality Management	4	26	38	3	71
2019	Mikalef et al.	Journal of Business Research	0	18	40	5	63
2019	Sun et al.	Tourism Management	1	12	36	9	58
2018	Dubey et al.	International Journal of Logistics Management	1	13	39	3	56
2018	Batista e Silva et al.	Tourism Management	2	19	28	4	53
2018	Coble et al.	Applied Economic Perspectives and Policy	7	19	21	4	51
2019	Polyakova et al.	International Journal of Economics and Business Administration	0	36	14	0	50

Table 4. Number of citations during 2018-2021 for the collected 377 references, excluding self-citations

Publication info			2018	2019	2020	2021	Total
Year	Authors	Journal Title	70	631	1750	342	2796
2018	Wirtz et al.	Journal of Service Management	5	38	127	19	189
2018	Li et al.	Tourism Management	4	42	90	22	158
2018	Lim et al.	Cities	2	37	46	9	94
2019	Allam and Dhunny	Cities	0	13	61	10	84
2019	Zhao et al.	International Journal of Hospitality Management	0	14	50	11	75
2019	Ardito et al.	Business Process Management Journal	0	13	48	9	70
2018	Amado et al.	European Research on Management and Business Economics	5	23	32	3	64
2018	Buhalis and Leung	International Journal of Hospitality Management	4	24	32	3	63
2019	Sun et al.	Tourism Management	1	12	33	7	53
2019	Mikalef et al.	Journal of Business Research	0	10	35	5	50
2018	Batista e Silva et al.	Tourism Management	2	18	26	4	50

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Word cloud of references list, titles of articles published, under the search limited to years 2018-2021 and for the retrieved 377 documents



CONTENT ANALYSIS OF THE MOST RELEVANT PUBLICATION (FULL PAPER ANALYSIS)

In this point, a content analysis of the 13 most relevant articles (shown in Table 3) was carried out. Although Table 3 includes 14 of the most cited articles for the shortened period of analysis, we had to resort to the content analysis to the 13 that were freely available to download, or for which we had access. For this purpose, the QDA Miner and WORSTAT software were used to extract and understand the information in the context of the documents analyzed. We are presenting in Figure 11 the clusters of the pertinent expressions or terms in the context of this study. It is particularly evident the importance that studies on consumer behavior, both for tourist products and other general services, assume; in which the focus is to understand their preferences and the satisfaction. The information typically used in these studies is big data produced by users result of their actions on the internet (such as web search data, online booking data, webpage visiting data, etc.) or data collected from several devices (such as GPS or mobile phones).

The analysis of all these data, customer ratings, and customers' perceptions allow to make forecasts and anticipate preferences and expectations of the consumers. The objective is to better understand customers to deliver improved customer service, make better decisions and increase the value of a business. The second group of works seeks to analyze how the large volume of data can improve the quality of life of a citizen, reaching more sustainable and inclusive cities and environments. To better understand these needs, the IoT ecosystems, the data stored based smart devices, and sensors are crucial. Thus, big data in this context contributes to the creation of information for stakeholders to make better decisions. The third group of studies highlights the importance of digital technologies and solutions in the domain

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Visualization of the content of the most relevant paper (N=13)

of Industry and organizations, to better used organizational resources regarding productivity gains. In this particular, aspects related to swift trust and humanitarian supply chains are highlighted.

CONCLUSION

Data Science and big data entered the mainstream academic vocabulary and the economy and business are no exception. The implementation range of big data in economics and finance has expanded from predicting financial distress to modeling efficient trading strategies, monitoring energy markets, managing the tourism industry, and socioeconomic process. This paper analyzed 377 publications on these topics in the scientific domain of firm performance, tourism, forecasting, customer satisfaction, and marketing, along with socio-economic monitoring, supply chain management, and smart cities development and interconnectivity. This growing literature encourages the contribution of big data techniques to the field of applied economics and business since these techniques can manage large volumes of data more practically without having rigid distribution assumptions.

There is a rapid growth in the use of data science and big data techniques. To demonstrate, decision tree methods in financial market strategies are used for the combination of multiple joint techniques

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

based on sequential explanatory strategies, which start with quantitative analysis and then end up using semi-structured primary data analysis. Over the years text mining advanced from the development in big data techniques, and machine learning strategies became popular especially in predicting the future of the tourism industry, customer satisfaction, socioeconomic process, and the development of smart cities. However, differences in the outcome of techniques indicate that there is no common methodology of data science and big data, and the development of techniques is still underway. As for data mining techniques, the subject area of execution is still in progress. Although major concepts, such as firm performance, potential in the tourism industry, forecasting, marketing strategies, smart cities, and interconnectivity with ecosystem have been reviewed, so far, there are many more topics in economics, business, econometrics, and finance that can be managed with the help of big data and data science techniques.

The current growth trends predict a large increase in the number of publications in these areas. The USA and UK made the most relevant contributions within this important field. followed by Italy, Germany, and Australia. Besides the contribution of technologically advanced countries, there are some important contributions to data science and big data from the emerging economies as well such as China, South Korea, South Africa, India, and Malaysia. Which could be a game-changer in the field of data science and big data as these economies are very big in term of human resources, consumer markets, emerging financial markets that can create a spike or accelerate the marketing techniques, financial management, consumer satisfaction based on big data and data science.

Excel, Stata, Matlab, and SPSS have been the most applied toolboxes when using data science in economics and business. But things are changing rapidly with a new generation of economists, which now embrace open-source software platforms like Python and R which are becoming the go-to platforms for many economists nowadays, mostly because these are free software's. Numerous economists are now working as data scientists in different fields. Similarly, many data scientists are working with economic and behavioral problems which were exclusively the domain of economists in the past. Every discipline is different but there are similarities as well in terms of problems solved, methods used, theoretical constructs, and tools applied. Data scientists must be able to acquire, understand, and analyze complex and unstructured data, interpret the meaning of the results, and communicate the results to relevant stakeholders. Hence, data science requires a unique combination of econometrics and statistics, programming skills, and domain knowledge. When combining data science to economics and business, where decision making is a constant, we get the perfect match in interpretation, explanation, and support for policy directions to be pursued.

There is an increasing trend in the application of data science models, methods, and techniques in both economics and business, still, huge opportunities and challenges are emerging. This chapter highlights some other research opportunities and avenues, mainly connected to big data existent opportunities. For example, corporate governance and image recognition, well-being and wealth and consumer's choices, opinions and reviews, game-based learning and education, trading robots and market price forecasts, traders blogs and prices, industry 5.0 emergent opportunities, high-frequency data analysis, all kinds of financial risk analysis and forecasting, development or growth forecasting, social corporate responsibility and performance, connecting air pollution to energy and environmental economics, business management decisions support, municipalities data and performance indicators, systems for decision support and policymakers policies decisions support, among many other subjects that we could elaborate here, which were detected as still lacking appropriate research linking data science with the economics and business fields. Furthermore, despite the rapid development in the evolution of the techniques, techniques performance and implementation area are still in need of improvement. The findings of this chapter

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

are, however, restricted to the economics and business areas and cannot be generalized to other fields of disciplines. We hope our findings will be useful to guide researchers and practitioners in economics and business areas to select the appropriate model after becoming familiarized with all these data science terminologies. We hope to encourage readers in the enlargement of their vision and data mining applications, with special emphasis on those connected and related to economics and business.

REFERENCES

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities (London, England)*, *89*, 80–91. doi:10.1016/j.cities.2019.01.032

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on big data in marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics*, *24*(1), 1–7. doi:10.1016/j.iedeen.2017.06.002

Ardito, L., Petruzzelli, A. M., Panniello, U., & Garavelli, A. C. (2019). Towards industry 4.0: Mapping digital technologies for supply chain management-marketing integration. *Business Process Management Journal*, *25*(2), 323–346. doi:10.1108/BPMJ-04-2017-0088

Bajracharya, K. (2018). *Data Science Glossary. Data Science.* Retrieved on 13 February 2021 from https://dimensionless.in/data-science-glossary/

Basdas, U., & Esen, M. F. (2021). *Review of Big Data Applications in Finance and Economics. In Handbook of Research on Engineering, Business, and Healthcare Applications of Data Science and Analytics*. IGI Global. doi:10.4018/978-1-7998-3053-5

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, *68*, 101–115. doi:10.1016/j.tourman.2018.02.020

Brooks, C., Hoepner, A. G. F., McMillan, D., Vivian, A., & Simen, C. W. (2019). Financial data science: The birth of a new financial research paradigm complementing econometrics? *European Journal of Finance*, *25*(17), 1627–1636. doi:10.1080/1351847X.2019.1662822

Buhalis, D., & Leung, R. (2018). Smart hospitality—Interconnectivity and interoperability towards an ecosystem. *International Journal of Hospitality Management*, *71*, 41–50. doi:10.1016/j.ijhm.2017.11.011

Cao, L. (2017). Data science: A comprehensive overview. *ACM Comput. Surv. 50*(3). doi:10.1145/3076253

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, *40*(1), 79–96. doi:10.1093/aepp/ppx056

Consoli, S., Reforgiato Recupero, D., & Saisana, M. (Eds.). (2021). Data Science for Economics and Finance. Methodologies and Applications. Springer International Publishing. doi:10.1007/978-3-030-66891-4

Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. doi:10.1080/10618600.2017.1384734

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dubey, R., Luo, Z., Gunasekaran, A., Akter, S., Hazen, B. T., & Douglas, M. A. (2018). Big data and predictive analytics in humanitarian supply chains: Enabling visibility and coordination in the presence of swift trust. *International Journal of Logistics Management*, *29*(2), 485–512. doi:10.1108/IJLM-02-2017-0039

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. doi:10.1016/j.ejor.2017.11.054

La Torre, M., Botes, V. L., Dumay, J., & Odendaal, E. (2019). Protecting a new achilles heel: The role of auditors within the practice of data protection. *Managerial Auditing Journal*. Advance online publication. doi:10.1108/MAJ-03-2018-1836

La Torre, M., Botes, V. L., Dumay, J., Rea, M. A., & Odendaal, E. (2018a). The fall and rise of intellectual capital accounting: New prospects from the big data revolution. *Meditari Accountancy Research*, *26*(3), 381–399. doi:10.1108/MEDAR-05-2018-0344

La Torre, M., Dumay, J., & Rea, M. A. (2018b). Breaching intellectual capital: Critical reflections on big data security. *Meditari Accountancy Research*, *26*(3), 463–482. doi:10.1108/MEDAR-06-2017-0154

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323. doi:10.1016/j.tourman.2018.03.009

Lim, C., Kim, K., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities (London, England)*, *82*, 86–99. doi:10.1016/j.cities.2018.04.011

McAfee, A., Brynjolfsson, E., & Davenport, T. H. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68. PMID:23074865

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261–276. doi:10.1016/j.jbusres.2019.01.044

Moon, K. S., & Kim, H. (2019). Performance of deep learning in prediction of stock market volatility. *Economic Computation and Economic Cybernetics Studies and Research*, *53*(2/2019), 77–92. doi:10.24818/18423264/53.2.19.05

Ndou, V., Secundo, G., Dumay, J., & Gjevori, E. (2018). Understanding intellectual capital disclosure in online media big data: An exploratory case study in a university. *Meditari Accountancy Research*, *26*(3), 499–530. doi:10.1108/MEDAR-03-2018-0302

Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S. S., Reuter, U., Gama, J., & Gandomi, A. H. (2020b). Data Science in Economics: Comprehensive Review of Advanced Machine Learning and Deep Learning Methods. *Mathematics*, *8*(10), 1799. doi:10.3390/math8101799

Nosratabadi, S., Mosavi, A., Keivani, R., Ardabili, S., & Aram, F. (2020a). State of the Art Survey of Deep Learning and Machine Learning Models for Smart Cities and Urban Sustainability. In A. Várkonyi-Kóczy (Ed.), *Engineering for Sustainable Future. INTER-ACADEMIA 2019. Lecture Notes in Networks and Systems* (Vol. 101). Springer. doi:10.1007/978-3-030-36841-8_22

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Polyakova, A. G., Loginov, M. P., Serebrennikova, A. I., & Thalassinos, E. I. (2019). Design of a socio-economic processes monitoring system based on network analysis and big data. *International Journal of Economics and Business Administration*, *7*(1), 130–139. doi:10.35808/ijeba/200

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking* (1st ed.). O'Reilly Media.

Sun, S., Wei, Y., Tsui, K., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1–10. doi:10.1016/j.tourman.2018.07.010

Taddy, M. (2019). *Business Data Science: Combining Machine Learning and Economics to Optimize* (1st ed.). Automate, and Accelerate Business Decisions.

Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems with Applications*, *143*, 113042. doi:10.1016/j.eswa.2019.113042

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, *29*(5), 907–931. doi:10.1108/JOSM-04-2018-0119

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111–121. doi:10.1016/j.ijhm.2018.03.017

KEY TERMS AND DEFINITIONS

Algorithms: A set of instructions that are followed to solve the problem.

Analytics: Systematic analysis of data or statistics.

Artificial Intelligence: Simulate human intelligence in machines that are programmed to think like humans and imitate their actions.

Big Data: Huge in volume or too large data and complex, in nature.

Cloud Computing: To deliver different services through the Internet. These resources include tools and applications like data storage, servers, databases, networking, and software

Data Mining: The process of finding differences, patterns, and correlations within large data sets to predict outcomes.

Deep Learning: It structures algorithms in layers to create an artificial neural network that can learn and make an intelligent decisions on its own.

Ensemble Learning: The process by which various models, such as classifiers or experts, are strategically generated and combined to solve a particular computational intelligence problem

Machine Learning: Use algorithms to parse data, learn from that data and make an informed decision based on what it has learned.

Optimization: Finding the best available.

Text Mining: Transforming unstructured text into structured data for easy analysis.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 27
A Historical Review of Immersive Storytelling Technologies:
New Uses of AI, Data Science, qnd User Experience in Virtual Worlds

Hector Puente Bienvenido
Universidad Complutense de Madrid, Spain

Borja Barinaga
Universidad Francisco de Vitoria Madrid, Spain

Jorge Mora-Fernandez
University of California, San Diego, USA

ABSTRACT

This chapter is focused on describing the history and the current relevance of user experience (UX) techniques that combine data science and AI in the research field of interactive and immersive storytelling, including virtual and augmented realities. It initially presents a brief history of interactive storytelling, video games, VR and AR, AI and data science, and the user experience (UX) techniques used in those areas. Later, the chapter describes the UX techniques in depth, using AI and data science that work best and are more useful for testing interactive media products, describing examples of its applications briefly. Finally, the chapter presents conclusions in relationship with utopias and dystopias regarding the future use of UX, AI, and data science in several areas such as edutainment, social media, media arts, and business, among others.

DOI: 10.4018/978-1-7998-6985-6.ch027

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

There has been a significant implementation of technology to manage scientific data during the last two decades. One consequence of the increased use of technology has led to diverse findings in science and high techs like VR systems, UX, and big data analytics, that have challenged and questioned our descriptions and representation of reality and our daily habits.

This chapter focuses on the material senses. Significant optimization and economization of virtual reality (VR) and augmented reality (AR) displays have served to question, explore, and research the communication and languages of the material senses to a greater extent so that they can be merged and immersed in different VR and AR products. This new reality has reinforced a more profound recreation of the physical senses. Advancements in applied data science, the increase of bandwidth for data transfer networks, mobile ram abilities and memories, massive media alternatives for displaying and collecting social information, and the implementation of programming systems have helped Artificial Intelligence (AI) to manage significant amounts of personal information. These improvements are mainly motivated by the desire to increase business and industry profits through the registration, analysis, and monetization of the collected data from users and data based on the human interactions with different high-tech interfaces (computer, mobiles, VR and AR displays, and tablets). Of course, we can use high technology for other purposes, such as education or improving human beings' wellness. However, business and industry are still the leitmotivs of most research and applications of data science and artificial intelligence.

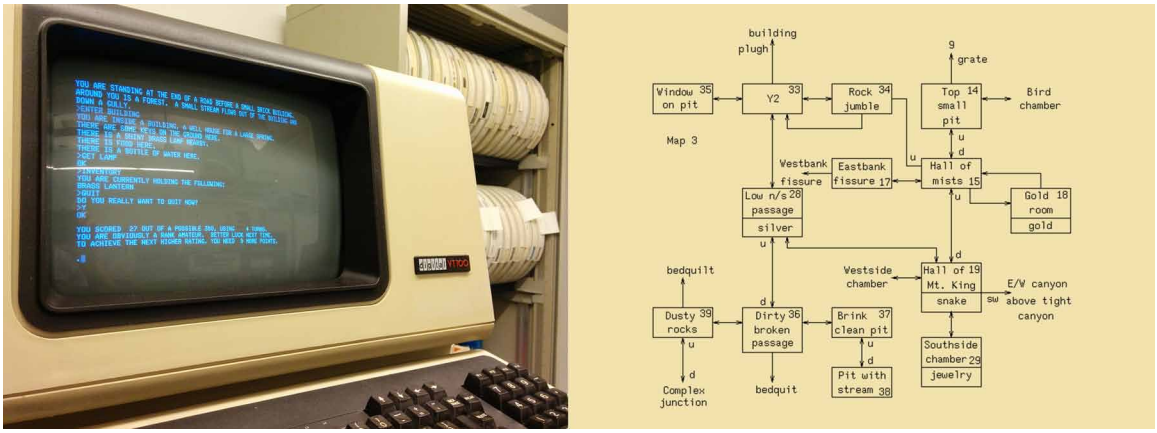
In this complex context of massive subjective and objective data and the ability of managing it with artificial intelligence software, companies are investigating the possibility of using data science, which gathers large amounts of subjective experiences, to collect objective responses from their users through the analysis artificial intelligence can provide.

HISTORICAL REVIEW OF INTERACTIVE, IMMERSIVE AND TRANSMEDIA STORYTELLING UX

Automated and interactive narration is usually linked to video games. However, it is not entirely true that its origin is only in computer games. Many of the characteristics of these narrative systems can be found in board games and, significantly, in the birth of role-playing games, invented by Gary Gigax (1938–2008) and Dave Arneson (1947–2009). The game *Dungeons & Dragons* (1974) was revolutionary, for it changed the paradigm of the traditional game. What is more, it established the way to elaborate a free and interactive story, lived in the first person, and based on rules. With the popularization of video game consoles and personal computers, due to the cheapness of electronic components, many young engineers became interested in programming fantasy-themed worlds that were mainly influenced by the work of J.R.R. Tolkien (1892–1973). Another critical factor to emerge was connecting computers in networks, such as through Arpanet (whose development began in the 1970s with the NCP protocol, *Network Control Program*, before the current TCP / IP, Transference Control Protocol and Internet Protocol). It was William Crowther, one of the programmers involved in the development of Arpanet, who programmed the first fully interactive, text-based fantasy story: *Colossal Cave Adventure* (1976) (Figure 1) (Barinaga, 2010). In 1977, the interactive story was polished and expanded by Don Woods, one of the fathers of hacker culture, who, as a student, discovered *Cave Adventure* on a Stanford University computer.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Left: Colossal Cave Adventure (1977) running on a PDP-11/34 and displayed on the VT100 console. Author: Autopilot. Right: a detail of a graph showing the structure of the world from the Colossal Cave Adventure Page (https://rickadams.org/adventure/index.html)



What was revolutionary about the system invented by Crowther was that it performed some very useful simplifications to produce the behavior of a purely artificial world. Colossal Cave Adventure's world has characters, space, and time – the essential elements for any narrative. Crowther conceived a world organized by spaces and connected by paths, much like the galleries of a cave. In each space, the player finds objects or characters, and the player can explore the world by picking up and putting down items (Barinaga, 2014). Notably, *Colossal Cave Adventure* has the fundamental characteristics of an interactive story, as defined by Crawford, which are structure, people, conflict, puzzles, choices, and spectacle (Crawford, 2005).

Significantly, the AI used in this program differed from the conventional one that was linked to the classical game studies, especially the game of chess, based on the research of Claude Shannon (1916–2001), used in emblematic projects like the *IBM Deep Blue* (1996) or *Deepmind/Google Alpha Go* (2015). The AI in the video game is much simpler but more effective from a narrative point of view. The Crowther structure simplifies reality without requiring complex AI and is based on simple state changes and automation. Despite its simplicity, the structure creates a rich experience, offering a great sense of freedom and a complex narrative. *Cave Adventure's* fundamental difference from other video games is that the AI is primarily used to achieve the feeling of spectacle that Crawford mentioned (Crawford, 2005). The program was designed to make the player's experience something memorable. Thus, the intention is not to simulate human decision making, but rather simulate plausible situations that give a sense of reality. Crowther's program proposed the following key aspects to simplify computing:

- Narrow paths and connections between each space or room that are justified by the narrative. This approach maintains the feeling of freedom while drastically limiting the possible options and facilitating programming.
- The appearance of an inventory that allows the taking and leaving elements. The changes of state caused by these interactions in the matrices serve as a trigger and conditioner of the adventure according to the algorithms of the general rules.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Writing recognition system and narrator simulation. The game was revolutionary for allowing a natural communication between player and program. When the program was not able to interpret an order, it responded narratively using a random system of phrases and a large dose of humor. This forced the player to seek other alternatives to express themselves or perform actions.

As previously described, since the proliferation of role games and the first computers, computational videogames, especially conversational adventure video games, caused the public to gain more interest in interactive narratives. The main goal became discovering how to combine immersive storytelling with different engaging interactive structures and levels of interactivity so the user experiences could improve, and people could feel more engaged in their narrative experiences and game dynamics.

Emergent forms of interactive cinematic arts exploring these subjects in interactive video games (e.g., virtual and augmented reality, among others) were presented for decades in international conferences and festivals all over the world.

Research groups such as the Interactive Cinema at MIT, publications like The Language of New Media by Lev Manovich of the Visual Arts Department and CRCA at UC San Diego, and the interactive documentaries displayed on museums by the Labyrinth Project led by Marsha Kinder at the University of Southern California, inspired interactive media academic programs. That is how the Interactive Digital Media Programs at the University of Southern California, Carnegie Mellon University, and the Georgia Institute of Technologies started. The last one was led by Janet Murray (1997), the author of the book *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* and who two decades later theoretically established the field (Murray, 2012).

In Europe, a significant exhibition and MIT publication entitled *Future Cinema* integrated the best research groups' findings and media artworks under the same roof at the ZKM, Center for Art and Media. The exhibition was curated by Peter Weibel and Jeffrey Shaw in Karlsruhe, Germany. In Spain, the Medialab Madrid was created around 2003 by Karin Ohlenschläger and Luis Rico, at the Centro Cultural Conde Duque. The center encouraged new media artists to research and create in the areas of interactive media and digital storytelling. *Via Intima* (Inside Track) was the first independent interactive transmedia project with an interactive exhibition, online, and interactive DVD. The project was developed at Medialab Madrid by Daktyl4 and led by Jorge Mora-Fernández and Borja Barinaga. Barinaga became one of the designers of the first academic videogame program in Spain. These creative and explorative spaces served as the first spaces for creation and testing of the emergent interactive digital storytelling arts and user experience (UX).

In terms of theoretical pioneers who persisted and consolidated research studies on interactive narratives, Marsha Kinder (1991, 2014), Michael Mateas (2004), Janet Murray (1997, 2012), Marie-Laure Ryan 2001), Lev Manovich (2005, 2001), Steve Anderson (2011), and Henry Jenkins have helped to establish the descriptions of some critical concepts regarding interactive storytelling and digital media. These concepts include, immersion, agency, and later transmedia. In Spain, in the area of interactive museums, Isidro Moreno (2002) from the Universidad Complutense de Madrid, director of the research group Museum I+D+C, has been an important pioneer. He is the author of the Interactive Museum of the Book of the Spanish National Library and author of the book *Musas y Nuevas Tecnologías*. Following Moreno's investigation lines, Jorge Mora-Fernandez, while researching with some of the mentioned American researchers at UCSD and USC, developed a model of analysis of interfaces to determine the most meaningful variables of interactive and transmedia narratives to study under UX to generate more interactivity, agency, and immersion. These models later evolved into a model of analysis for transme-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

dia projects. In terms of the interactive and new forms of the book, the last years have been relevant to see the formation of the UNESCO New Forms of the Book, chaired by Alexis Weedon, UK, and the research developed by Deglaucy Jorge Teixeira, in Brazil, as well as the international conferences in digital storytelling, ICIDS.

Regarding the digital culture evolution from interactive media storytelling to transmedia storytelling, Marsha Kinder initially used the transmedia concept (1991, 2014) referring to transmedia intertextuality. However, Henry Jenkins is the one who popularized the transmedia narrative term. Jenkins (2011) gave the following description of the term in his blog: "transmedia storytelling represents a process where integral elements of a fiction get dispersed systematically across multiple delivery channels for the purpose of creating a unified and coordinated entertainment experience. Ideally, each medium makes its own unique contribution to the unfolding of the story."

Previously, Jenkins had described in his book, *Convergence Culture* (2006), the concept of convergence as a paradigm for thinking about the present moment of the media change. He defined the change by the stratification, diversification, and interconnection of the media in a converged way that influences the decisions of media producers, politicians, and citizens on the production and consumption of the digital culture. More recently, and because of the latest communicative and cultural practices, Jenkins evolved the convergence concept to become the transmedia term (2011):

Transmedia, used by itself, simply means "across media." Transmedia, at this level, is one way of talking about convergence as a set of cultural practices. We might also think about transmedia branding, transmedia performance, transmedia ritual, transmedia play, transmedia activism, and transmedia spectacle, as other logics.

In this regard, the transmedia narrative generates a universe of linear and circular possibilities where the stories start and finish with feedback to themselves and experimented one at a time. The linearity of the interactivity in this regard is understood as the coherent, communicative, immersive, and multilevel flow: aesthetic, narrative, and emotional,

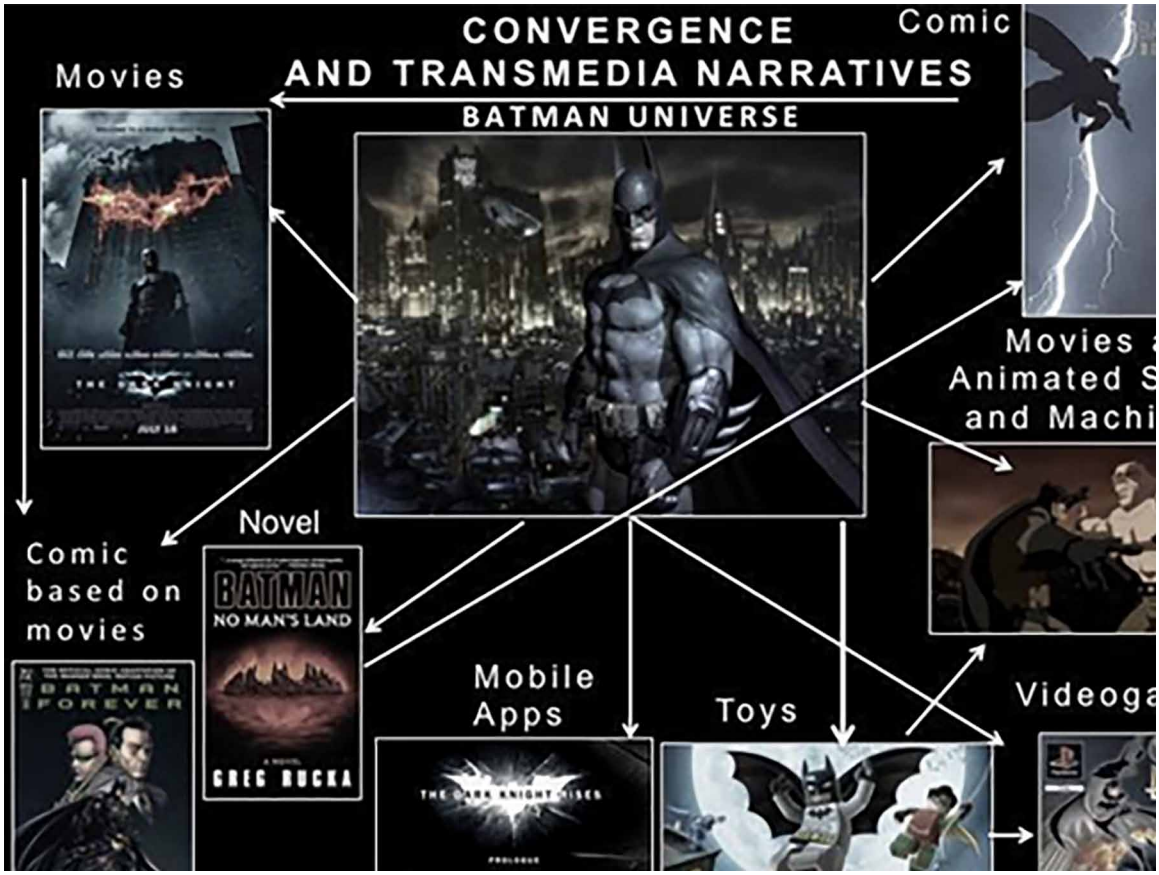
Mora-Fernández (2013). Interactive linearity is a concept that has been developed in detail through time after the digital narrative and cultural evolution from hypermedia interactive narratives Mora-Fernández (2009). See Table 1 for the variables included on the model of analysis to interactive transmedia narratives Mora-Fernández (2019a, 2019b). See Figure 2 for the story flow across-media.

AI and Data Science Use in UX Interactive, Immersive, and Transmedia Storytelling

During the last years, conferences in interactive digital storytelling have been proliferating. An example of these conferences includes the International Conference on Interactive Digital Storytelling (ICIDS). At these events interactive digital storytellers in the areas of cinematic arts, videogames, and virtual reality exchange their project experiences of experimenting with interactive and transmedia digital storytelling systems. Some notable collections and research projects in this interdisciplinary arena are by Nunes, Oakley, and Nisi (2017); Marcus and Rosenzweig (2020); Nack ad Gordon (2016); Schoenau-Fog et al. (2015); Long (2000); Barbic et al. (2017); Mora-Fernandez (2019a, 2019b); Ioannidou (2013); and Parody (2011).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. This image visually summarizes the influence and complementary story between media in one of Batman's Universes (from comic, to movie, to animation, to games to book)



It is essential to say that the integration of aesthetic and narrative variables on a human-computer interface is so vast, almost as vast as the human imagination, that some AI used on UX only focuses on very specific aspects, as in shown the next chapters. Therefore, it is presented here as a model that has taken into consideration all the data science needed to perform fundamental research on interactive storytelling. Among others, the most meaningful aesthetic and narrative variables to observe in order to implement the interface product qualities in interactivity, immersion, and agency have been taken into consideration (see Table 1). It is interesting to see how Unity is allowing us to observe more variables for dynamic design adjustment (DDA) purposes. Soon, AI will collect all this data and compare the results between the interactive digital storytelling variables to finally respond to the fundamental question that motivates any UX test in any interactive media product: What are the most effective and attractive aesthetic? And what are the most compelling narrative elements in the interface that motivate and interact in different ways?

Because of the massive number of variables, the UX in interactive digital storytelling (IDS), UX data science, and AI is very simplified. Most of the time the IDS UX tests rely on traditional user experience questionnaires, such as the one used in the research *Connecting the Dots: Quantifying the Narrative Experience in Interactive Media* by Hannesson et al., Schoenau-Fog et al. (Eds.) (2015). However, some

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Model for analysis of hypermedia interfaces to observe and register the contained aesthetics, narrative elements, type of interactions and ethical values (adapted from Mora-Fernández, 2012)

1. Name and description of the interfaces and the conjunction of hypermedia expressions 1.1. Identifiable denomination of the hypermedia interface. Each different interface that appears on the hypermedia product should be numbered, named and observed under this descriptive model to obtain a deep analysis 2. General characteristics of the interface and detailed description of the multimedia characteristics of the expressions that can allow interaction with any of the narrative elements 2.1. Software: Group of expressions and technological tools that are used for the relationship and generation of natural and virtual interactions A) Of iconic intermediation B) Symbolic C) Combination of A & B D) Natural mimetic • Opened or virtual reality • Semi-opened or simulators of virtual reality F) Convergent G) Pull or push interfaces H) Static or dynamic Interfaces I) Mute or sound interfaces J) Smart interfaces K) The iteration 2.2. Types of image or perceptive representations A) Still image 1) Photo-mimetic 2) Photo-infographic 3) Info graphic B) Still image with sounded image C) Image in movement 1) Cine-mimetic 2) Cine-infographic 3) Cine-mimetic and infographic D) Visual image in movement with or without sounded image E) Audiovisual image F) Sounded image G) Sounded image with or without visual image or extraterritorial images 2.3. Hardware: Group of physical expressions A) Of intermediation B) Natural-mimetic 2.4. Typographic description • Size of letter • Style of font or type • Characteristics or effects of the letter • Color of the letter 2.5. Iconic description 2.6. Symbolic description 3. Features of the characters represented on the interface and general description of the potential interactions with the characters 3.1. Character or avatar of 1st, 2d or 3rd Person	3.2. Physical characteristics • Sex • Age • Height and weight • Hair, eyes and skin colors • Pose • Corporal appearance and customs • Morphological defects • Hereditary aspects 3.3. Sociological characteristics • Economic status • Employment • Type of education • Life and family relationships • Religion • Race, nationality • Function in his community • Political tendencies 3.4. Psychological characteristics • Sexual and moral life • Personal ambitions and motivations • Frustrations, main conflicts • Temper: angry, tolerant, pessimistic, optimistic, etc. • Vital attitude: complacent, combative, surrendered • Insecurities: obsessions, inhibitions, superstitions • Extroverted, introverted, well balanced • Capacities, aptitudes, languages • Qualities: imagination, criteria, taste, equilibrium • Intellectual coefficient: high, regular, low 4. Interactional aspects of the character and type of interaction available: selective, transformative or constructive 5. Values or spiritual principles and unscrupulous values that available to activate through the interaction with the narrative characteristics of the characters. Values and unscrupulous values that appear potentially related with the interaction developed 5.1. Ethical values 5.2. Unscrupulous values 6. Characteristics of the actions represented on the interface and general description of the potential interactions with the actions 6.1. Type of structure 6.2. Secondary theme or subplot 6.3. Changing hierarchy 6.4. Changing hierarchy •Relationships between main and secondary actions •Real relationships between main actions •Real relationships between secondary actions •Simulated relationships between main and secondary actions •Annulated between main and secondary actions 7. Aspects of the interactional actions and type of interaction available: selective, transformative or constructive	8. Values or spiritual principles and unscrupulous values that available to activate through the interaction with the narrative characteristics of the actions. Values and unscrupulous values that appear potentially related with the interaction developed 8.1. Ethical values 8.2. Unscrupulous values 9.Characterists of the spaces represented on the interface and general description of the potential interactions with the spaces 9.1. Natural, constructed, mimetic-natural or mimetic-info graphic 9.2. Senses implied in the spatial perception: view, ear and/or touch 9.3. Implicit space and/or explicit 9.4. 2D/3D or 4D space 9.5. Perspective: size, scale, position and point of views 9.6. Focus or defocus 9.7. Illumination and color temperature 9.8. Props 9.9. Protagonist space and/or hyperspace 9.10. Absent space or suggested space 9.11. Selection space with representation: coincident or different 9.12. Hyperspace 10. Aspects of the interactional spaces and type of interaction available: selective, transformative or constructive 11. Values or spiritual principles and unscrupulous values that available to activate through the interaction with the narrative characteristics of the spaces. Values and unscrupulous values that appear potentially related with the interaction developed 11.1. Ethical values 11.2. Unscrupulous values 12. Characteristics of the time represented in the interface and general description of the potential interactions with the time 12.1. Order: flashback, flashforward, meta-retrospective or meta-prospective 12.2. Duration: pure diegesis, impure diegesis, open or close 12.3. Frequency: repetitive sequence or singular multiple 12.4. Temporal localization: past, present, future, changing or inexistent 12.5. Iteration 13. Aspects of the interactional times and type of interaction available: selective, transformative or constructive 14. Values or spiritual principles and unscrupulous values that available to activate through the interaction with the narrative characteristics of the times. Values and unscrupulous values that appear potentially related with the interaction developed 14.1. Ethical values 14.2. Unscrupulous values

projects developed and programmed their tools to complement the user experience questionnaires. Some interesting ones found on Nunes, Oakley, and Nisi (2017), and Schoenau-Fog et al. (Eds.) (2015) are:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Centre for Game by Aamri and Greuter, Schoenau-Fog et al. (Eds.) (2015).
- *Mise-en-scène: Playful Interactive Mechanics to Enhance Children's Digital Books,* Design Research, RMIT University, Melbourne, Australia. This research uses "two instruments Reading Motivation Scale (IRMS), and direct observations during the three school reading sessions… ...The *Motivation for Reading Questionnaire* (MRQ) was designed by Wigfield and Guthrie, and it measures 11 dimensions, including reading efficacy, challenge, curiosity, reading involvement, importance, recognition, grades, social competition, compliance, and reading work avoidance."
- *Connecting the Dots: Quantifying the Narrative Experience in Interactive Media Story Immersion in a Gesture-Based Audio-Only Game* by Wu and Rank, Schoenau-Fog et al. (Eds.) (2015), from Drexel University, Philadelphia, PA, USA. The researchers used an audio-based UX. They "compare useful designs of audio feedback that are responsive to hand gestures to both encourage user participation and maintain immersion in gesture-controlled audio-only environments. For the experimental study, two designs for responsive audio feedback were implemented in two game versions, as described below. We set up a between-subjects experimental design in order to compare the two game versions and their effect on the evaluation of the game, including its immersive qualities (…). The method is based on three moments: Experiment setup: hand motion input, processing, and audio output. The hardware setup for the game includes using a Leap Motion Controller for hand tracking and data input, a laptop for data recognition and processing of the game, and noise-cancelling stereo headphones for audio output."

- *Generating Side Quests from Building Blocks* by Tomas Hromada, Martin Cerny (B), Michal Bıda, and Cyril Brom, Schoenau-Fog et al. (Eds.) (2015) from the Department of Software and Computer Science Education, Charles University in Prague, Prague, Czech Republic. Their research developed "multiple automated quests generating systems were described in the literature. The GrailGM system uses a forward chaining inference engine to recombine quest elements given by a designer to form quests with multiple solutions and to guide the player's progression through the game. Petri Nets have also been proposed as models for generating quests. One of our primary goals was to keep the system simple and controllable while expanding the expressive power of template-based systems. To achieve both, we decided to split quests into blocks. Each block represents a short section of a quest (e.g., killing a monster to obtain an item). The blocks are arranged into a directed acyclic graph (DAG) where edges connect each block to its possible successors. The DAG may be represented explicitly (as in our implementation) or given implicitly by conditions. Blocks that have no incoming edges are called start blocks, and those with no outgoing edges are end blocks."
- *Evaluation of Yasmine's Adventures: Exploring the Socio-Cultural Potential of Location Aware Multimedia Stories* by Dionisio, Barreto, Nisi, Nunes, Hanna, Herlo, and Schubert, from Madeira-ITI, University of Madeira, and the Design Research Lab, Berlin University of the Art, Berlin N. Nunes et al. (2017). Their research entitled. In the research there is a combination of using Nvivo software. "The questionnaire and interview data were then analyzed using a grounded theory affinity analysis with the Nvivo software tools support. Researchers conducted a bottom-up data analysis reviewing and iterative process in three stages: open coding (quotes and high-level categories), affinity diagrams (relationships between categories), and theme organization (most frequent themes and description)."

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

In terms of interactive digital storytelling using a collection of specific data in combination with AI in UX AI Mobile and VR narrative experiences, some interesting examples appeared in *Interactive Storytelling* (Nunes et al., 2017):

- Developed by Kampa and its research Authoring Concepts and Tools for Interactive Digital Storytelling in the Field of Mobile Augmented Reality, from Hochschule RheinMain, Wiesbaden, Germany SPIRIT appears like an exciting UX tool. "In SPIRIT, a Storytelling XML (STARML) dialect has been derived from ARML by adding authoring-friendly XML-tags, with the focus on location-based content description for IDS. Further, a plot engine has been developed that interprets the STARML content structure, processes user interactions, and triggers AR video and other media. This engine uses conditions for planning. Authoring AR content has been made accessible for programmers by systems like EDoS and ComposAR. For non-programmers, TaleBlazer offers a visual script language like Scratch. An authoring concept is observing the authors for authoring tool development (...). The SPIRIT authoring tool StoryPlaceAR was developed out of the need to author media content outdoors in an ad-hoc manner (…). The system "immersive authoring for tangible AR" transforms 'What You See Is What You Get' authoring concept into 'What You Feel Is What You Get'. VideoTestAR enables prototypical AR video production and automatically authors the STARML content structure for immediate experience testing. In authoring for IDS, after content creation, the story structure can be altered by filling in content into an existing story structure. For this, several working steps are performed on the same STARML file and folder system. This process creates a bottleneck of collaboration and communication in an interdisciplinary team. In conclusion, this bottleneck exists for the authoring of XML structures for IDS systems."
- Developed by Bala, Nisi, and Nunes in their research Evaluating User Experience in 360o Storytelling Through Analytics at Madeira-ITI, University of Madeira, they developed their evaluation VR analytics for narrative structures "Based on the need for evaluation of complex narrative structures in VR storytelling, we designed IVRUX, a VR analytics platform for 360o video. This tool, intended to be used in an academic context, offers a support system to facilitate communication of creators and/or researchers with participants. The system is composed of three components: an online dashboard for researchers (built using the Angular framework) and a mobile VR application for the participants (built using Unity), connected through a REST API backend (Node.js and MongoDB). The system is designed so that the researcher does not need to have any expert skills (e.g., coding) to evaluate content."

In conclusion, IDS UX testing that can provide useful information for data science and AI, but there is a long way to go. However, software by iMotions, which integrates eye-tracking, EEG, and emotional facial recognition, has emerged as a promising tool for interactive communications and storytelling research. Some exciting applications of interactive storytelling on recovery from addictions can be found in research by Mora-Fernandez et al. (2020) entitled *iMotions' Automatic Facial Recognition & Text-Based Content Analysis of Basic Emotions & Empathy in the Application of the Interactive Neurocommunicative Technique LNCBT* (Line & Numbered Concordant Basic Text).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The video game Pong (Atari, 1972) is a perfect example of UX principles with very limited technological possibilities. This game put Atari at the forefront of the new entertainment industry. Exhibition in the Tekniska Museet, Stockholm, Sweden (16 September 2014). Photography by Diderot

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

BRIEF HISTORY OF UX APPLIED TO VIDEOGAMES

The design systems used in video games undergo constant optimization and evolution towards a fully adaptive user experience. They are applied both in the gaming medium and transferred to the most varied experiences by hybridizing the digital and physical world (Sequeiros & Puente, 2018, 2019). The innovations in design systems respond to a growing audience's demands. Therefore, one made up of people with unique characteristics, such as desires and interactions. Video games as cultural devices have been adapted to users through tools that allow the monitoring and analyzing of a player's performance or preference, enabling designers to adjust and tune the user's experience (Puente & Sequeiros, 2019a, 2019b). Dynamic design adjustment (DDA) not only allows the game to be altered in real-time based on the decisions and behavior of the player, but it also enhances the enjoyment and social inclusion of the game without neglecting the capacities and possible limitations of the users. DDA thereby follows the design justice approach (Crews & Zavotka, 2006; Constanza-Chock, 2020).

Traditionally a closed medium with predetermined systems, video games and their social contexts have been intertwined and merged with today's reality. Although past attempts, such as *Tempest* (Atari, 1981) and *God Hand* (Clover Studio, 2006), were limited by technology, a design approach that allows the game experience to be adapted to the player's interests has emerged. As it has been demonstrated, usability is a critical element in interactive systems (Maguire, 2001). For example, Missura and Gärtner (2009) have emphasized that the ideal video game should adapt its difficulty dynamically based on the player's performance.

From its origins, the field of videogames has been a cutting-edge platform for UX studies. Interfaces, gameplay, and narrative maintain, in most cases, a critical dependency to improve the player experience. After the burst of the first video game consoles into homes, such as Ralph Baer's *Odyssey* (Magnavox, 1972), the personal computers' current architecture was established: a motherboard, a sound card, and a graphics card. From that moment on, the interface development related to image and sound materialized in revolutionary projects such as the human-computer interfaces for the *Apple Lisa* (Mackintosh 1983), linked to Steve Wozniak and Steve Jobs (1955–2011) who were both ex-collaborators from Atari. (Kent, 2001)

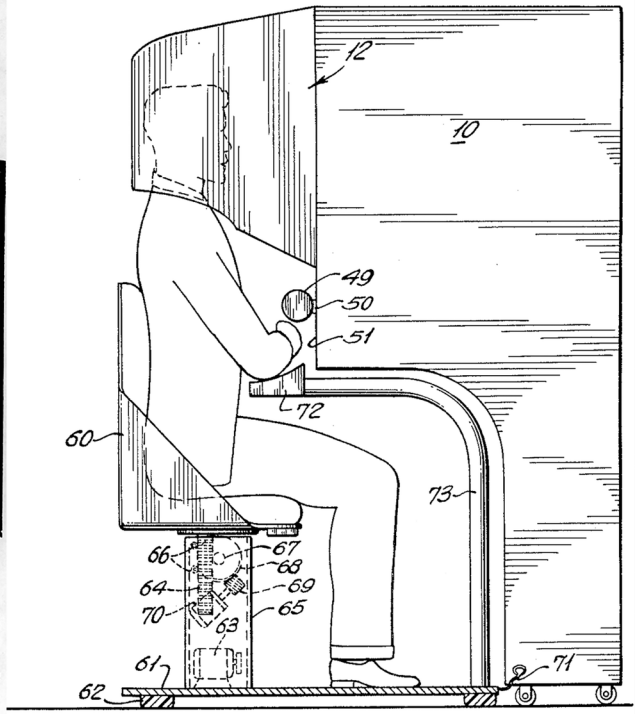
In the eighties, Atari did not know what kind of industry it was creating. However, they did know the importance of narrative immersion and the experience of the players in their games. This is the reason they created research and development departments that thought about the future of interfaces and new options for human-computer interaction as a key to success (Kent, 2001). Today, some of these historical interfaces are part of the collections of prestigious museums, as is the case of *Pong* (1972; Figure 3), which is in the MOMA and the Science Museum of London, among others.

Unlike other fields, in the video game industry the importance of UX is preponderant. Games are based on the player's experience, and the activity is not compulsory but voluntary. If the experience is not rich and exciting, and the mechanics, gameplay, and narrative are not in line with the interface's quality, the player will stop playing, and the product will be a failure (Schell, 2008).

In this sense, from the origins of the main mechanics, it can be said that a massive incursion of artificial intelligence in the development of video games and in the improvement of the UX has been experienced, highlighting three key aspects:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Sensorama advertising poster and illustration of Morton Heilig's Sensorama device, a precursor to later virtual reality systems (Filed Jan 10, 1961). Author: Morton Heilig. Figure 5 of U.S. Patent #3050870 (via https://patft.uspto.gov/)



- With the increase in complexity of video games, automated and intelligent help and management systems have emerged that improve the quality of the gameplay and minimize the importance of secondary tasks not related to the experience or gameplay (Millington, 2006).
- Data analysis and information flow methods have been improved to facilitate testing and design, even from the game engines themselves.
- Increases in processing speed allows for the management of the intelligent automatic behaviors of a significantly greater number of non-player characters at the same point of the game. (Millington, 2006)

MOST RELEVANT MILESTONES IN VR, AND AR USER EXPERIENCE

Virtual reality has more than fifty years of existence now. However, in the last few years VR technology and industrial uses have been growing dramatically. The advances in electronics, energy (more efficient batteries), and computer technology have made it possible to realize a dream that began to develop from the multi-sensory simulator *Sensorama* (1960), created by Morton Heilig (Figure 4).

Sometimes the concept of VR is fuzzy, due to the large number of uses in the context of virtual spaces generated by 3D computer technology. Nevertheless, VR is mainly related to the simulation of a space by computer graphics to allow the user, intermediated through a hardware interface, to interact

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. HTC Vive (2016) vs VPL Research DataSuit (Circa 1989). HTC Executive Director of Marketing Jeff Gattis wears the HTC RE Vive (3 March 2015). Author: Maurizio Pesce. VPL Research DataSuit, a full-body outfit with sensors for measuring the movement of arms, legs, and trunk developed circa 1989. Displayed at the Nissho Iwai showroom in Tokyo (4 October 1999). Author: Dave Pape

with things in a similar way that he or she would in the real world. The VR experience is immersive by definition and tries to be multi-sensory. (Fuch, 1992; Gigante, 1993).

The first head-mounted display (HMD) of VR was *The Sword of Damocles*, constructed by Ivan Sutherland (Sutherland, 1965). The system was able to track the orientation of the head to update the virtual scenario.

In the seventies there were relevant systems with multiple uses, such as the *GROPE* (1971). This was a human-force-feedback developed at the University of North Carolina. Today, this type of system combined with our advances in data transference could allow, for example, doctors to practice surgery remotely, systems like the *Da Vinci Surgical System* (Intuitive Surgical, 2000), present in approximately five thousand hospitals in 2020 (Singer 2010). Another early VR invention was *VIDEOPLACE* (1975) by Myron Krueger (Krueger et al., 1985). This hardware was the origin of all the movement reconnaissance systems used in later video game hardware, like *Microsoft Kinect* (2010) and *Nintendo Wii* (2006).

In the eighties, new interests were found in the military field for the use of virtual spaces. In 1981, Atari collaborated with the US Army to develop a training simulator for the gunners of the M2 Bradley Tank, inspired by the successful arcade game *Battlezone* (1981; Barinaga 2016). In 1982, the US Air Force, with the Armstrong Medical Research Laboratories, developed the *VCASS* (Visually Coupled Airborne System Simulator), an HMD system to control the pathways of missiles and targets (Mazuryk & Gervautz, 1999).

In the mid-eighties, the first VR commercial devices appeared. Examples include the *Fake Space Labs BOOM* (Binocular-Omni-Orientational Monitor, 1989) and the *HMD Data Glove* (1985) by the VPL company (Figure 5).

But in the end, all these devices, and the advances made in the 1990s, were too costly and bulky to reach the entertainment industry. Moreover, they remained as anecdotic, military, or scientific experiments with expensive hardware. In 2016, two companies launched to the market affordable and compatible new VR glasses systems: the *HTC Vive* (HTC) (Figure 5) and the Oculus Rift (Oculus VR). The videogame industry has embraced this technology as an up-and-coming field for the new generation of games. Furthermore, VR glasses have started to become a convenient interface for design and 3D modeling.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Although augmented reality shares similarities with VR, it is a visualization of reality with virtual content that augments the standard visual information. Thus, it enriches rather than replaces the real world (Bryson, 1992).

In 1990, Tom Caudell coined the concept of *Augmented Reality* while working on a project to simplify the view of wire schematics for airplanes at the Boeing company, and he developed with his colleague, David Mizell, the idea for special glasses that show information in real-time (Carmigniani et al., 2011). However, the advances in AR remained in a scientific or military context until 1999. The technology became more accessible when Nara Institute of Science and Technology's Hirokazu Kato released the *ARToolKit*, an open-source tool kit for any OS platform to capture video and insert, in real-time, virtual information. Since the AR modules are present in the OS of smartphones, the uses of the technology have been exponentially growing. In 2013, Google launched the *Google Lens*, and in 2016 Microsoft developed the *Holo Lens* for Windows Mixed Reality. Nintendo revolutionized the game industry with the AR game *Pokémon GO* (2016; Figure 6). Today, AR content can be viewed with smartphones and is used habitually in commercials, sports, and video games.

Increasing amounts of research using AI and data science in AI and VR have been developed during the last years. Some exciting publications that have collected interesting research are Jung and Tom Dieck (2018), Peddie (2017), as well as some proceedings from the EuroVR International Conference, Barbic et al. (2017), among others.

AI AND DATA SCIENCE IN STORYTELLING AND VIDEO GAME DESIGN

Artificial intelligence and data science have been linked to the development of commercial video games since their origins. Character simulation, non-player character (NPC) movements and actions, algorithms, and predictability, among others, play a fundamental role in gameplay experiences (Bekkes et al, 2012; Nantes et al, 2008) (data science applications are especially visible in areas like serious games).

The first game to partially use an AI dates back to the 1950s. This is the case with Nim. His AI implemented combinatorial game theory using binary logic operations. Two decades later, games such as *Pong* (1972), *Space Invaders* (1978), and *Galaxian* (1979) would already include simple AIs that incorporated preset patterns based on examining the input (action) of the human user. A few years later *Tempest* (1981) appeared. It was the first video game to use player performance parameters to optimize their gaming experience and achieve a balanced level of challenge. Thus, data science applied to video games was born and would revolutionize the field of game balancing adjustment with optimized matchmaking, generation of rewards, and item-drop mechanics.

Today, dynamic design adjustment (DDA) has been proven successful in various media such as video games and interactive experiences (Maguire, 2001). Significantly, DDA has been shown to improve the performance and the engagement of the users (Andrade et al., 2006; Xue et al., 2017). Adjusting NPCs' behavior to make them less aggressive, adapting the number of enemies in an environment and their health or damage values, and increasing the recovery speed of the health bar of players are some alternatives and design decisions that have been applied in studies (Chapman & Hunicke, 2004). DDA has also been as applied to favor the user experience in matchmaking (Aghdaie et al., 2007). However, the use of dynamic design adjustment is not limited to mechanics. The narrative can also be altered by adapting the narration to the player's profile. Graphics engines like Unity already have their own tool to implement dynamic design and make this data more accessible for development studios.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Pokémon GO fever in Japan (12 October 2019)

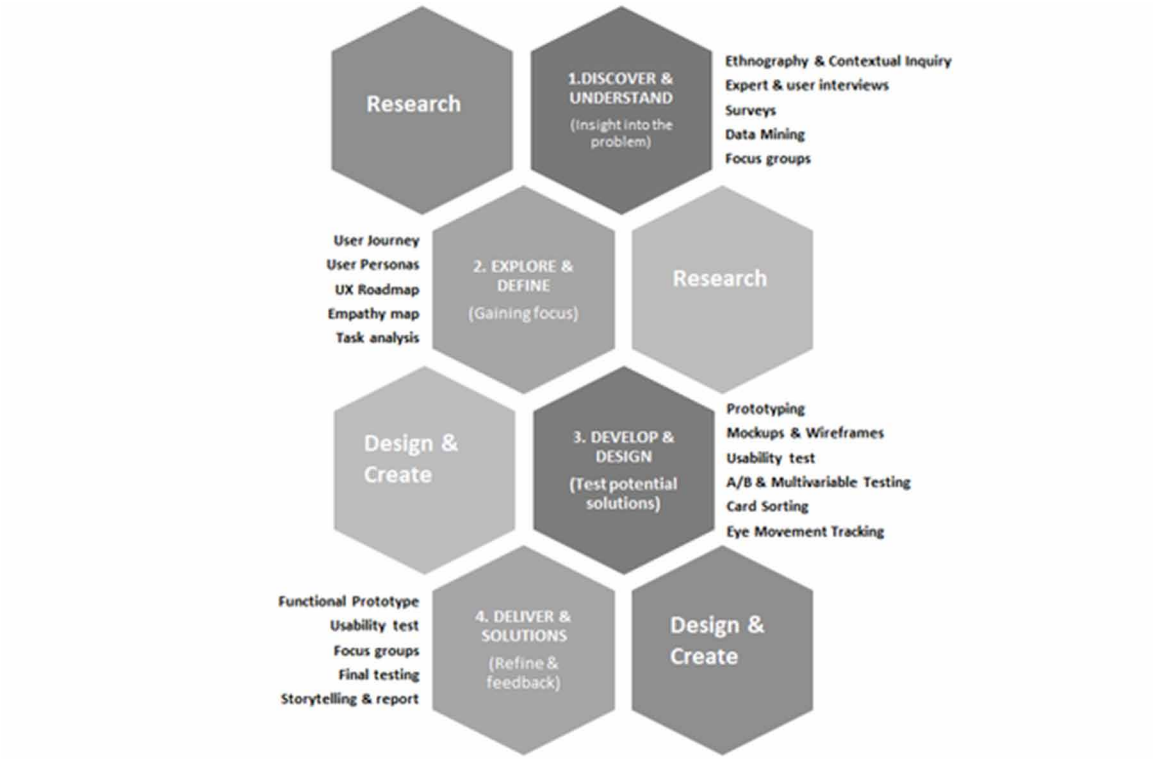
Corporations also have DDA systems based on machine learning. This is the case with IBM Watson. Through algorithms, Watson undergoes constant learning and adaptation, allowing for the development of unique systems for each player and individualization of the user's experience.

This approach opposes the traditional vision of the video game as a fixed and pre-designed experience without adaptive measures to the behavior and interest of the player. The traditional structure of videogames can be observed in the GDDs (*Game Design Document*) of different video games, such as *Prince of Persia* (Broderbund, 1993), *Grim Fandango* (Lucas Arts, 1998) and *Planescape: Torment* (Black Isle Studios, 2009).

The main point is not only located on the adaptation of video games to the player, but on how this interaction is carried out. With Niantic and its video game *Pokémon GO* (2016) (Figure 6) as the maximum exponent, the trends of video games have been observed to leave the digital medium and merge with reality. Pervasive video games have been integrated, taking the mechanisms and dynamics of prior and 'classical' games to other formats such as AR. The augmented reality format is acquiring new titles, such as *Five Nights at Freddy's AR: Special Delivery* (Illumix, 2019), *Shin Megami Tensei Liberation Dx2* (Atlus & Sega, 2018), and the currently in development *The Witcher: Monster Slayer* (Spokko, TBA). There is an interest in developing and improving AR-based experiences.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Different UX techniques that are being used in the design of immersive experiences and services shown as steps (inspired by Design Council's Double Diamond)



Storytelling and gaming systems based on data science and AI have permeated their environment (Mateas, 2003). We can currently observe multiple gamified and data-driven experiences in museum tours, digital narratives, and dating apps. For instance, the Prado Museum in Madrid has an interactive room (room number 39) where VR is used to interact with art pieces, which again demonstrates the importance of this new technology.

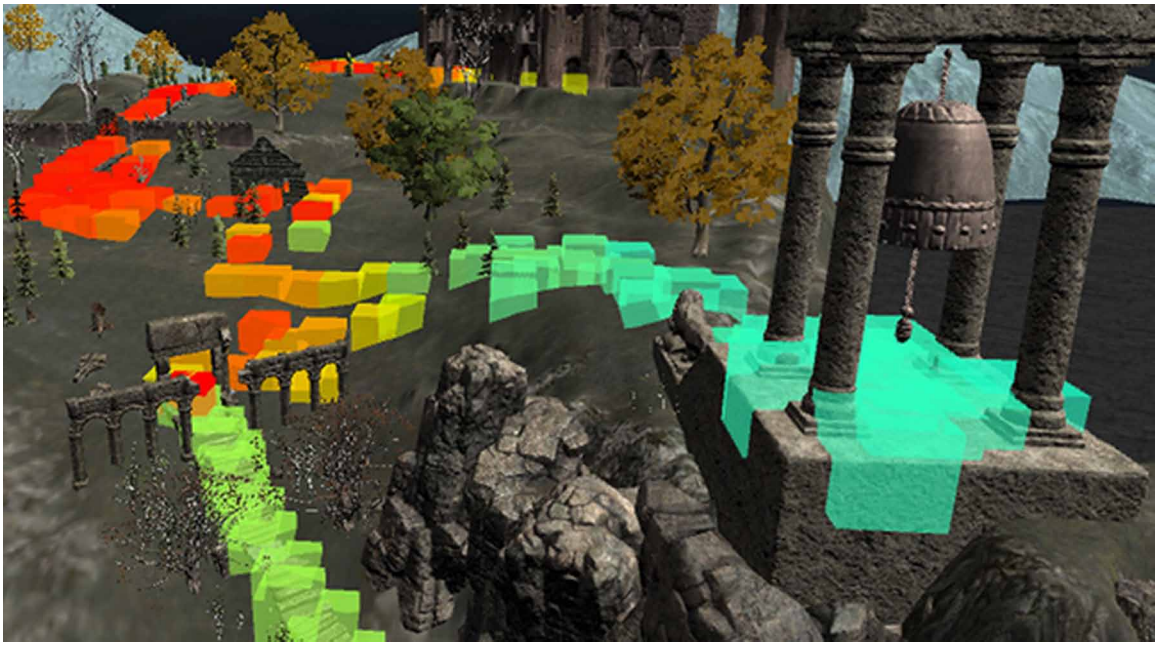
Thus, data science and artificial intelligence are critical pieces in the search for memorable experiences that consist of living an experience rather than consuming a product or story. The aforementioned methods are progressively approaching an increasingly immersive experience based on data analysis, which could be called Peak Immersive Experience. Using embedded metrics in devices, such as AR or VR, video game designers are encouraged to take advantage of the hardware paradigm change and establish DDA as a must-have from now on.

METHODOLOGIES TO IMPLEMENT UX INTERACTIVE AND IMMERSIVE STORYTELLING VR AND AR

After observing the great field that data science and artificial intelligence have to cover, this chapter focuses on using AI and applied data science in user experience techniques.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Unity Analytics heatmap (showing players' areas of interaction)
Source: Unity Analytics Documentation, 2021



First, a mapping of the different UX techniques (organized in design phases) that are currently being applied in the field of media storytelling must be explored.

Figure 7 identifies the different UX techniques that are being used in the design of immersive experiences and services. Inspired by the Design Council's Double Diamond, we have carried out an analysis of the research techniques most commonly used in the different phases of the design process. Notable is how an increasing influence of AI and applied data science in user experience techniques can be observed.

In this progressive integration of UX, AI, and data science, differentiation can be made between two relevant types of processes in terms of design systems and methodologies.

1. How do AI and data science collect information?
2. What narrative and aesthetic variables are currently considered most relevant or significant?

Although there are numerous examples of data analytics and tracking algorithms used in audiovisual devices, the current trend is to take advantage of the information collection options integrated into the development engines themselves (such as Unity or Unreal).

Any pattern of play, behavior, event, or interaction is recorded in a database, which can now be handled very intuitively via a data dashboard. This ability allows to collect and cross all the information on variables of interest for our analysis.

The narrative and aesthetic variables that are currently considered the most relevant or significant are the following: visual and special effects, generation of emotions (Fernández et al., 2021), quality in level design, and story (cutscenes, narrative interactions, etc.).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

- quality, diversity, and customization of avatars,
- number of interactions (by click or collision) with game objects,
- level of concentration, emotions (very relevant in sensory storytelling) and attention on environment and user interface design elements (measured through biometrics, interactions and heat maps; see Figure 8),
- percentage of space to map explored (including percentage of completion or secrets discovered),
- the possibility of "skipping" or avoiding the story (cutscenes, textboxes) in narrative experiences (variables of great interest are the total and relative time of reading and listening to the story, the percentage of scrolling made to the textboxes, the number of skips performed, the number of interactions with other non-player characters),
- relative weight of the narrative (in time and relevance) on the gameplay and game experience (measured by game test and questionnaire, engine tracking events, etc.) and
- multidimensional measure of creativity (Jackson et al., 2012).

Engines like Unity already have their own tool to implement dynamic design and make it more accessible for development studios. Thus, GameTune and Unity Analytics are positioned as a dissociative element of DDA and high-budget video games, as Unity is one of the most used engines for video game development. The more variables that can be identified and contracted, the closer developers will be to having a complete map of what works and what does not in UX applied to interactive media storytelling.

IMPLEMENTING USERS' INTERACTIONS AND IMMERSIVE EXPERIENCES IN DIGITAL STORYTELLING

Reflections of What Works and What Does Not

The progressive rise of digital mediations and UX design techniques has led to a productive orientation that finds in the creation and sale of immersive and designed experiences an important niche of commercial, socio-cultural, and political exploitation. In this sense, authors such as Pine and Gilmore (2011) describe the emergence of a new type of economy that began at the end of the 20th century and is based on the production of highly personalized interactive experiences and digital storytelling. Producing memorable and captivating experiences to be sold and enjoyed is a trend that has always been latent, particularly in the entertainment field, but in recent decades this trend has exploded and grown significantly.

As early as the 1990s, Myron Krueger (1991) predicted that the information society (Castells, 1996) was going to be replaced by one in which experience would be the commodity with value. He also affirmed that, at that time, it was already visible that there were many ways of buying experiences and that access to them was already considered as wealth. For Krueger, virtual reality, in any of the forms in which it has ended up being established (e.g., videogames and their platforms), would be the paradigmatic example of this "experience society."

Accordingly, we have been flooded by immersive products that sell us digital storytelling experiences in packages for all types of profiles. Most importantly, this trend is also reaching other aspects of social reality such as interpersonal relationships (social networks and dating apps), politics (participatory democracy or digital activism), and leisure (escape rooms, interactive, performative, and immersive amusement parks, and e-sports). Nevertheless, this is a trend that goes beyond mere entertainment. For example, it

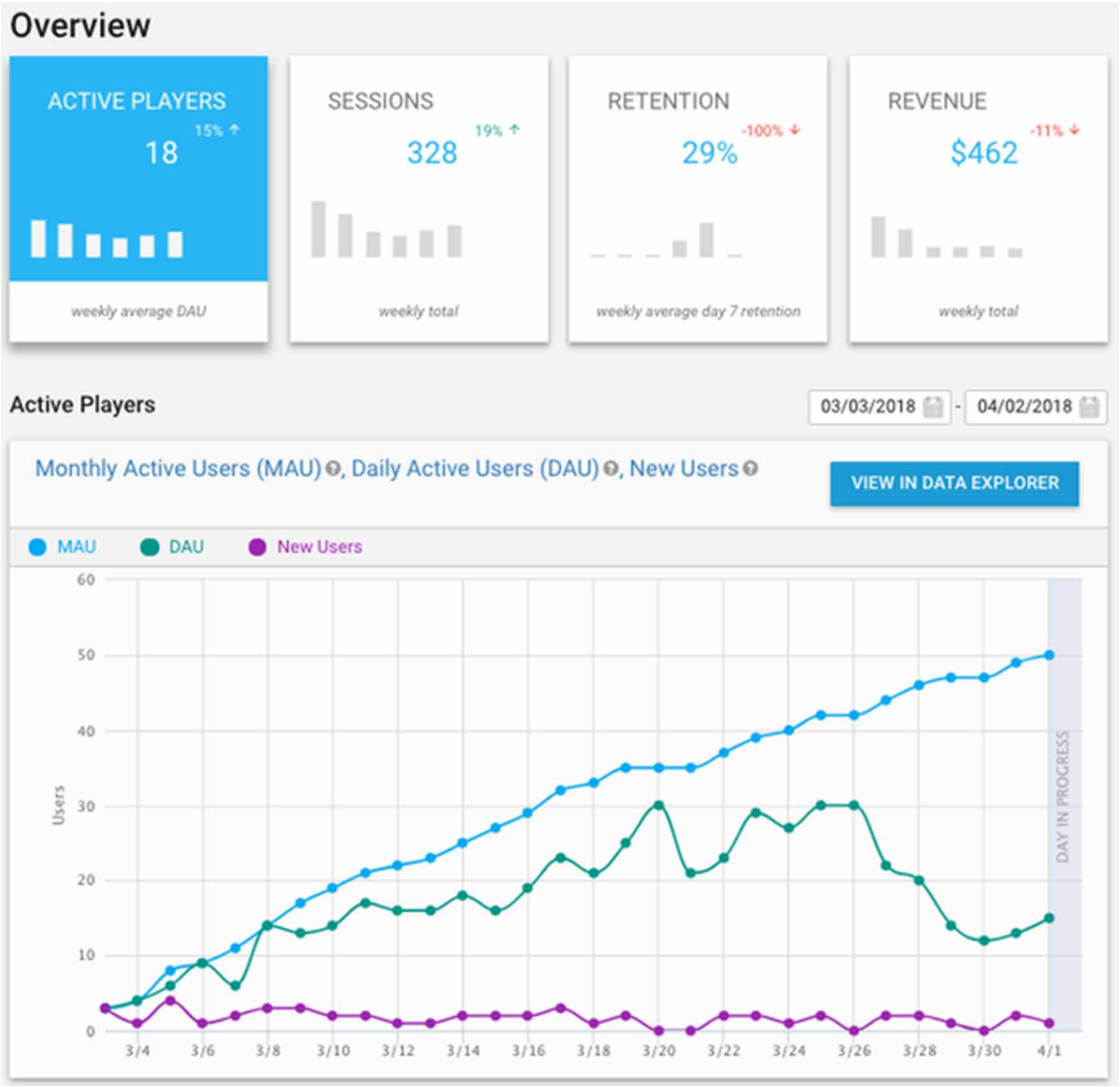
EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Facial emotional recognition during the participant interactions where each line represents different types of basic emotions (adapted from iMotions)

is not particularly difficult to find fields that are being structured, to a lesser or greater degree, in terms of experience. Tourism and cultural heritage are areas that have been especially permeated by the staging of experiences and by the sale of experience packages (sightseeing tours, dramatized wine tastings).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Unity Analytics interface showing players' key performance indicators (KPIs) (adapted from Unity Analytics Documentation, 2021)



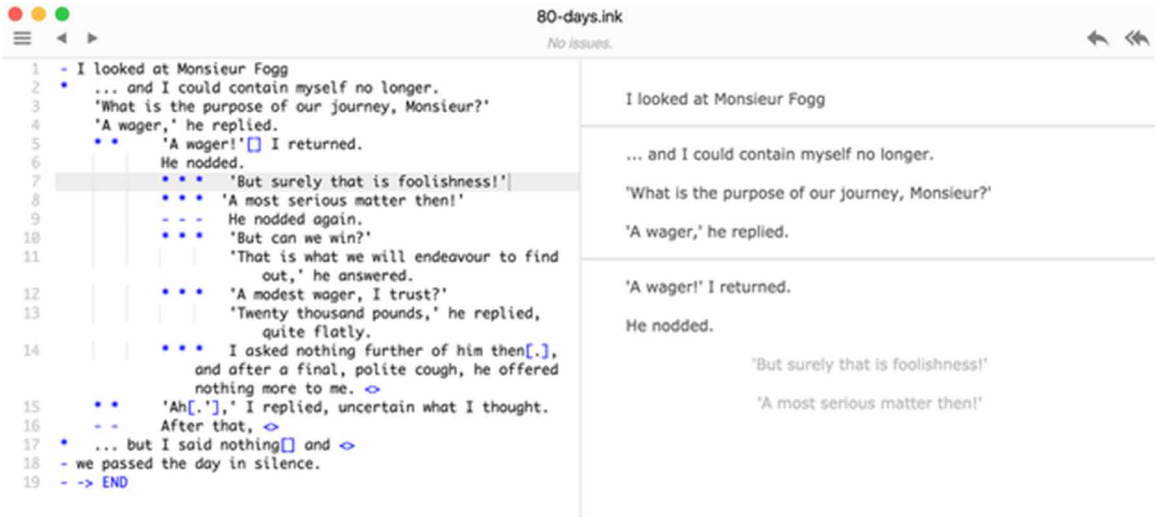
Storytelling and collecting or processing data in real-time have a significant influence on the afore mentioned experiences (Fernández & Puente-Bienvenido, 2015), and today there are numerous examples of great interest tools.

Some Implementations Suggestions in UX Methods and Software for Implementing Interactive and Immersive Storytelling

There are numerous tools in the audiovisual entertainment industry that enable the implementation and the study of highly immersive and interactive storytelling. Some of the tools are very popular, as in the

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Inky editor interface (adapted from Inkle Studios, 2021)



case of Twine (an open-source tool for telling interactive, nonlinear stories) and iMotions (research software for access to emotional, cognitive, and behavioral data; see Figure 9).

However, the most potent innovations have appeared thanks to their integration in game engines (which allow informational tracking in real-time and readjust the experience continuously). In the following list three main tools that we consider to have the most significant potential in interactive and immersive storytelling due to their efficiency, accessibility, and synergies (taking advantage of data science and AI in game engines) are presented.

1. **GameTune:** Based on machine learning and AI, GameTune automatically optimizes elements such as progression speed, play patterns, difficulty, and game narratives through A/B testing. This UX tool continuously learns, adjusts, and iterates to deliver the best possible gaming experience, which means more engaged players and a better gameplay experience. GameTune can be used for various problems where the goal is to select the best alternative from A, B, C, (…) for each user. The best alternative is the one that maximizes engagement, player experience, immersive storytelling, conversion, and a custom game-specific reward.
2. **Unity Analytics:** Unity Analytics is a simple but powerful data platform that provides UX data in a video game (mostly related to a player's in-game behavior). Unity collects data by activating events when someone uses an application or game. Unity automatically dispatches a set of core events to collect richer and more in-depth information about how users play a game or experience storytelling. In addition to behavioral events for tracking user experience and player behavior, there are fully customizable events available whose names, purposes, and parameters can be defined as needed. Finally, it is also possible to use heat maps that are significantly useful in UX analysis.
3. **Inky:** Inky is an online tool for writing complex interactive stories that allow narrative designers to play and test their stories in real-time. It uses a powerful narrative scripting language that is

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

primarily designed for professional game development, though Inky can also be used to write and share choice-based interactive fiction.

CONCLUSION

The following conclusions are either based on the present emergent experiences or are imagined, but close to realistic tendencies. The conclusions covered in this chapter offer final observations in the areas of edutainment, social media, business, media arts, as well as other areas.

Ethical Consideration of UX Interactive and Immersive Storytelling in VR and AR

In the collective imagination of technology users, there are ambivalent discourses about how technical devices facilitate many areas of life and, in parallel, pose numerous risks and uncertainties. These are the so-called benefits and technological development counterparts (Sanmartín & Megías 2020; Sequeiros & Puente, 2019).

Fears and resistances towards technology do not occur uniformly, and certain areas, processes, and practices are perceived in very different ways. Perceptions can range from a partial fear of the implementation of a new technology in the field of interpersonal relationships, for reasons such as dehumanization, to an explicit desire to automate or digitize tasks that do not require much emotional involvement, such as health diagnosis, art, business, or media and audiovisual entertainment.

As a final compilation, we would like to go through a series of examples applied in different academic interest fields.

It is an excellent time for edutainment to develop given the need of more online education. To make full use and benefit of the transmedia and videogame entertainment industries, edutainment will require public institutions to invest in the new generation of interactive entertainment media to develop educational content.

The future of entertainment products linked to digital technology lies mainly in the video game industry. Within the frame of videogames, many possibilities coexist in gameplay, such as using different hardware and software interfaces. However, there is no doubt that gradually VR and AR are becoming relevant tools for game design. The involvement of VR and AR requires new ways of narrating and new possibilities to explore original audiovisual formats. UX-focused studios are already a vital part of the production phases of the new generation of video games. The massive data collection and the increased processing speeds of the new generations of processors allow designers to understand the user' behaviors within simulated or hybridized worlds much more effectively. Despite this advancement, the complexity of human behavior and the myriad factors to consider in the mathematical analysis of the information obtained do not allow absolute independence from game designers and their creative intuition. Nevertheless, the use of UX techniques dramatically simplifies and improves the work and results in game design.

The use of video games as a format for education is not without controversy. Fundamentally, if games become related to education, we are already introducing a political and ideological factor that is difficult to reconcile with an aseptic view of the issue. Merely talking about *serious games* (Clark Abt, 1970) as opposed to *video games* as entertainment introduces an idea linked to the medieval Christian tradition

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

by which games are related to vagrancy, perversion, and crime. If a game is something serious, it will cease to be a game. With videogames, ideas and culture are transmitted in the same way as with any other audiovisual medium. To obtain edutainment, a right balance between serious contents and their presentation must be applied.

What is true is that because videogames have a robust mathematical logic base in their structure, they are beneficial in promoting scientific thinking. The possibility of generating experiences in simulated environments is fundamental not only for leisure, but it is also crucial in simulation. Simulators are, habitually, used to train professionals of different fields, such as pilots and scientists. Moreover, in many cases, the line between the simulator and video game is minimal. For example, many vehicle simulators are also marketed as video games.

On the other extreme, the border between sports and simulators is more complicated. In recent years, new forms of technological and digital sports have been born. Furthermore, these new sportive practices are linked to the immersion in new artificial or hybridized realities. There are already athletes who race professionally in virtual car races with simulators. Drone pilots use FPV glasses (First Person View) with real and augmented information incorporated. Furthermore, there are drone competitions within simulators. The boundaries between reality, simulation, and game are challenging to establish. The game and the ludic experience can arise in any context and any space: real, virtual, or imagined.

The video game is undoubtedly an ideological weapon and is a transmitter of profound values of the societies in which we live. However, the ideological use of games is not something new. In the Greek world's initiation rites and in some of its representations like the Olympic Games, there was a fundamental political and social component for social cohesion. Nevertheless, if we combine the video game's immersive capabilities with massive data management, the anonymity of young users, and, in some cases, severe problems of privacy rights, a very interesting conjunction that can be used unethically by both companies and governments emerges.

Edutainment needs to develop on an ethical and moral base that can communicate important cultural and scientific contents while keeping the game focused on entertainment. UX techniques can help develop better interactive media products that enhance the development of human intellect and spirit and support data to complement accessibility and disability challenges. Edutainment can serve the professors and students to develop better empathic environments where knowledge is built together.

Human relations and social media are fields where digitization, data science, and UX have developed more intensely. Human relationships are increasingly digitally mediated, from applications for dating to applications for sharing social, political, and daily content.

One of the main impacts of AI and data science is the standardization and dehumanization of job categories and skills. This influence of AI has made challenges to new job searchers navigating the online selection processes developed by websites such as LinkedIn. New ways of developing resumes and CVs are being developed to enable AI to at least select the job searcher profile. The exactitude of the words and the lack of synonyms and imagination concerning the interconnection between certain skills makes job search apps very robotic and leaves out potentially interesting profiles that do not correspond to rigid categories.

A series of gamified tools with interesting interactive narratives and storytelling have emerged in this type of context. Using algorithms and data intelligence, game designers and data scientists try to adapt and customize the user experience as much as possible. Thus, technological devices commit the user to immersive stories, suitable matches, challenges, and achievements to unlock, but at what cost? The management of private information, a loss of anonymity, the traceability of data, and the segmentation

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of socio-demographic data for marketing purposes are some of the tolls to users pay in exchange for immediacy, permanent connection, access, and personalization.

Ambivalent desires for digitization and customization arise in areas where greater penetration is desired, coexist with others where the public shows certain fears for their potential risks. Thus, utopian discourses of accessibility, immediacy, integration, and the breaking of limits coexist in parallel with other more dystopian visions associated with surveillance, control, deprivation, and inequality (of access and management). Thereby, desires and expectations regarding the progressive introduction of AI and data science in our social interactions are traversed by situations of socio-demographic variability and some groups are made vulnerable to being exposed to the less attractive effects of technological developments' progressive penetration (Sanmartín & Megías 2020). For example, people with a lower educational level perceive more significant risks and vulnerabilities associated with technological development, because they are the ones who are most exposed to its risks. Conversely, highly educated groups perceive technology and UX in a more positive, libertarian, and innocuous way (technophilic or utopian vision).

DEDICATORY

In Memoriam of our beloved friend Jorge Mora Fernández (1976-2021), Interactive Media Researcher (University of California, San Diego, Arthur C. Clarke Center in Human Imagination), colleague, storyteller, and navigator of the Virtual Seas.

ACKNOWLEDGMENT

Thanks to Arthur C. Clarke Center for Human Imagination, University of California San Diego, USA; the Laboratory of Digital Culture & Hypermedia Museography-Research Group Museum I+D+C, Universidad Complutense de Madrid,Spain; and the Emoji Studies Research Group, Digital Narrative and Video Game Studies UFV Research Group, Universidad Francisco de Vitoria, Madrid, Spain, Laboratory of Digital Culture & Hypermedia Museography-Research Group Museum I+D+C, Universidad Complutense de Madrid, Madrid, Spain.

REFERENCES

Aghdaie, N., Kolen, J., Mattar, M., Sardari, M., Xue, S., & Atif-Uz Zaman, K. (2018). *Multiplayer Video Game Matchmaking Optimization*. United States Patent 9993735.

Anderson, S. (2011). *Technologies of History*. Dartmouth College Press.

Andrade, G., Ramalho, G., Gomes, A. S., & Corruble, V. (2006). Dynamic Game Balancing: An Evaluation of User Satisfaction. In *Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference* (p. 3–8). Menlo Park, CA: AAAI Press.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bakkes, S. C. J., Spronck, P. & van Lankveld, G. (2012). Player behavioural modelling for video games. *Entertainment Computing, 3*, 71-79.

Barbic, J., D'Cruz, M., Latoschik, M. E., Slater, M., & Bourdot, P. (2017). Virtual Reality and Augmented Reality. In *14th EuroVR International Conference, EuroVR 2017 Laval*, *France*, December 12–14, 2017 *Proceedings.* Springer. 10.1007/978-3-319-72323-5

Barinaga, B. (2010). *Juego. Historia, teoría y práctica del diseño de videojuegos*. Alesia Games.

Barinaga, B. (2014). Digital distopy: Videogame. In Artecnología Augmented Knowledge and Accessibility. Complutense.

Barinaga, B. (2016). The video game and the game tradition.Education and shadows of the digital society. *Opción*, *32*, 9.

Bryson, S. (1992). Measurement and calibration of static distortion of position data from 3D trackers. *SPIE Proceedings Vol. 1669: Stereoscopic Displays and Applications III*.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, *51*(1), 341–377. doi:10.100711042-010-0660-6

Castells, M. (1996). La era de la información. Economía, sociedad y cultura. Siglo XXI.

Chapman, V., & Hunicke, R. (2004). *AI for Dynamic Difficulty Adjustment in Games*. Academic Press.

Constanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds we Need*. MIT Press. doi:10.7551/mitpress/12255.001.0001

Crawford, C. (2005). *On Interactive Storytelling*. New Riders.

Crews, D. E., & Zavotka, S. (2006). Aging, Disability, and Frailty: Implications for Universal Design. *Physiol Anthropol.*, *25*(1), 113–118. doi:10.2114/jpa2.25.113 PMID:16617216

Cruz-Neira, C. (1993). Virtual reality overview. *SIGGRAPH 93 International Conference on Computer Graphics and Interactive Techniques*.

Fernández, M., Puente, H., & Sequeiros, C. (2021). *Nuevas formas de comunicación a través de la narrativa en tiempo real: producción virtual, videojuegos y otras experiencias interactivas.* Congreso Internacional Comunicación y Pensamiento 2021. https://comunicacionypensamiento.org/ponencia/nuevas-formas-de-comunicacion-a-traves-de-la-narrativa-en-tiempo-real-produccion-virtual-video-juegos-y-otras-experiencias-interactivas/

Fernández, M., & Puente-Bienvenido, H. (2015). Universos fantásticos de inspiración lovecraftiana en videojuegos survival horror. Un estudio de caso de PT (Silent Hills). *Brumal. Revista de Investigación sobre lo Fantástico/Brumal. Research Journal on the Fantastic.*, *3*(1), 95–118.

Fuchs, H., & Bishop, G. (1992). *Research Directions in Virtual Environments*. University of North Carolina at Chapel Hill.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gigante, M. A. (1993). Virtual reality: definitions, history and applications. *Virtual Real. Syst*., 3–14. . doi:10.1016/B978-0-12-227748-1.50009-3

Ioannidou. (2013). Adapting Superhero Comics for the Big Screen: Subculture for the Masses. *Adaptation*, *6*(2), 230–238. doi:10.1093/adaptation/apt004

Jackson, L.A., Witt, E., Games, A., Fitzgerald, H., Von Eye, A. & Zao, Y. (2011). Information Technological use and Creativity: Findings from The Children and Technological Project. *Computers in Human Behaviour, 28*(2), 370-376.

Jenkins, H. (2006). *Convergence Culture*. New York University Press.

Jenkins, H. (2007). *Transmedia Storytelling 101.* http://henryjenkins.org/2007/03/transmedia_storytelling_101.html

Jenkins, H. (2011). *Transmedia 202: Further Reflections.* http://henryjenkins.org/2011/08/defining_transmedia_further_re.html

Jenkins, H. (2013). *"Wish You Were Here": Imaginative Mobilities and Disembodied Intimacy in Postcards (Part Two).* http://henryjenkins.org/2013/02/what-transmedia-producers-need-toknow-about-comics-an-interview-with-tyler-weaver-part-two.html

Jung, T., & Dieck, T. (2018). *Augmented Reality and Virtual Reality. Empowering Human, Place and Business*. Springer. doi:10.1007/978-3-319-64027-3

Kent, S. (2001). *The Ultimate History of Video Games: From Pong to Pokemon—The Story Behind the Craze That Touched Our Lives and Changed the World*. Crown.

Kinder, M. (1991). *Playing with Power in Movies, Television and Video Games*. University of California Press.

Kinder, M. (2014). *Transmedia Frictions: The Digital, the Arts, and the Humanities*. University of California Press.

Krueger, M. (1991). *Realidad artificial 2*. Addison-Wesley Profesional.

Krueger, M. W., Gionfriddo, T., & Hinrichsen, K. (1985). Videoplace an artificial reality. *Proceedings of the ACM SIGCHI Bulletin*, 16. 10.1145/317456.317463

Long, G. (2007). *Transmedia Storytelling. Business, Aesthetics and Production at the Jim Henson Company* (Master Thesis). Massachusetts Institute of Technology.

Maguire, M. (2001). Methods to support human-centred design. *Human-Computer Studies*, *55*, 587–634.

Manovich, L. (2005). *Soft Cinema. The Language of New Media*. M.I.T. Press.

Mateas, M. (2003). *Expressive AI Games and AI, In proceedings of Level UP.* Digital Games Research Conference, Utrecht.

Mateas, M., & Murray, J. (2004). A preliminary poetics for interactive drama & games, & from game story to cyberdrama. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First Person: New Media as Story, Performance & Game* (pp. 2–33). The MIT Press.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mazuryk, T., & Gervautz, M. (1999). *History, Applications, Technology and Future*. Institute of Computer Graphics, Vienna University of Technology.

Millington, I. (2006). Artificial Intelligence for Games. Elsevier, Morgan Kaufmann Publishers.

Missura, O., & Gartner, T. (2009). *Player modelling for Intelligent Difficulty Adjustment*. Discovery Science, 12th International Conference.

Mora-Fernández, J.I. (2009). *La interfaz hipermedia: el paradigma de la comunicación interactiva. Modelos para implementar la inmersión juvenil en multimedia interactivos culturales. (Videojuegos, cine, realidad aumentada, museos y web).* SGAE. Fundación Autor. Ediciones Autor.

Mora-Fernandez, J. I. (2013). Artecnología en cine interactivo al unas categorías, interfaces, estructuras narrativas, emociones e investigaciones. In *ArTecnologia. Arte, Tecnologia e linguagens Midiáticas*. Buqui.

Mora-Fernandez, J. I. (2019a). Extended analysis of the transmedia narrative universe of Nolan's Dark Knight: integrating original authorships, media convergences & interactive linear narratives. In Reflections in Social Sciences and other Topics. Guayaquil: CIDE.

Mora-Fernandez, J. I. (2019b). Concepts and Models of Analysis of Interactive and Transmedia Narratives: A Batman's Universe Case Study. In *Information Technology and Systems ICITS 2019, AISC 918* (pp. 929–943). Springer. doi:10.1007/978-3-030-11890-7_87

Mora-Fernandez, J. I. (2020). IMotions' Automatic Facial Recognition & Text-Based Content Analysis of Basic Emotions & Empathy in the Application of the Interactive Neurocommunicative Technique LNCBT (Line & Numbered Concordant Basic Text). In Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Communication, Organization and Work. Springer. https://doi.org/10.1007/978-3-030-49907-5_5.

Moreno Sánchez, I. (2002). *Musas y Nuevas Tecnologías: El relato hipermedia*. Paidós Comunicación.

Murray, J. (1997). *Hamlet On the Holodeck: the Future of Narrative in Cyberspace*. Free Press.

Murray, J. (2012). *Inventing The Medium. Principles of Interaction, Design as a Cultural Practice*. The MIT Press.

Nack, F., & Gordon, A. S. (2016). Interactive Storytelling. ICIDS 2016, LNCS 10045, 61–72. DOI: 6 doi:10.1007/978-3-319-48279-8

Nantes, A., Brown, R., & Maire, F. (2008). A Framework for Semi- Automatic Testing of Video Games. In *Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference*. Association for the Advancement of Artificial Intelligence.

Nunes, O., & Nisi, V. (2017). Interactive Storytelling. *ICIDS 2017, LNCS 10690*, 372–375. doi:10.1007/978-3-319-71027-3_46

Parody, C. E. (2011). *A Theory of the Transmedia Franchise Character* (PhD thesis). University of Liverpool.

Piddle, J. (2017). *Augmented Reality. Where We Will Live*. Springer. doi:10.1007/978-3-319-54502-8

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pine, J., & Gilmore, J. H. (2011). *The Experience Economy*. Harvard Business Review Press.

Proietti, M. (2019). Experimental test of local observer independence. *Science Advances*, *5*(9). Advance online publication. doi:10.1126ciadv.aaw9832

Puente-Bienvenido, H. & Sequeiros, C. (2019a). A Sociological Look at Gaming Software: Erving Goffman's Dramaturgy in Video Games. *Revista Española de Investigaciones Sociológicas*, *166*, 135-152. doi:10.5477/cis/reis.166.135

Puente-Bienvenido, H., & Sequeiros Bruna, C. (2019b) Goffman y los videojuegos: Una aproximación sociológica desde la perspectiva dramatúrgica a los dispositivos videolúdicos. *Revista Española de Sociología*, *28*(2). https://recyt.fecyt.es/index.php/res/article/view/68948

Rico, L., & Ohlenschläger, K. (2005). *Banquete 05 – Communication in Evolution*. Museo Conde Duque.

Roth, C., & Koenitz, H. (2016). Evaluating the User Experience of Interactive Digital Narrative. In *Proceedings of the 1st International Workshop on Multimedia Alternate Realities*. Association for Computing Machinery. 10.1145/2983298.2983302

Ryan, M. L. (2001). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. Johns Hopkins University Press.

Sanmartín, A. y Megías, I. (2020). *Jóvenes, futuro y expectativa tecnológica*. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad. Doi:10.5281/zenodo.3629108

Schell, J. (2008). The Art of Game Design: A book of lenses. *Elsevier CRC Press*.

Sequeiros, C., & Puente-Bienvenido, H. (2018). Sociología y ci-fi: repensando los procesos de cambio social desde la ciencia ficción. *Nudos. Sociología, Teoría y didáctica de la literature, 2*, 4-16. doi:10.24197/nrtstdl.1.2018.4-16

Sequeiros, C., & Puente-Bienvenido, H. (2019). Cambio social, tecnología y ciencia-ficción. *Sociología y Tecnociencia*, *9*(2), 115–138.

Singer, E. (2010). The slow rise of the robot surgeon. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/news/418141/the-slow-rise-of-the-robot-surgeon/

Sutherland, I. E. (1965). *The Ultimate Display. Multimedia: From Wagner to Virtual Reality*. Norton.

Wilber, K. (1985). *The Holographic Paradigm and other paradox, exploring the leading edge of science*. New Science Library.

Xue, S., Wu, M., Kolen, J., Aghdaie, N., & Zaman, K. A. (2017). Dynamic Difficulty Adjustment for Maximized Engagement in Digital Games. *WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web*.

KEY TERMS AND DEFINITIONS

Game Design: Techniques and processes used to design the rules and contents of a videogame.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Game Design Document: Document with all the design information of a videogame. It´s the main document for designers, programmers, and artists in a videogame project.

Game Engine: The core software for a game to properly run. It allows the design, creation, and operation of a videogame.

Gameplay: The game experience that emerges from the interaction between the game design and the player.

Heatmap: A graphical representation of data in the form of a map in which colors represent data values.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 28
Artificial Intelligences Are Subsets of Human Beings, Mainly in Fictions and Films: Challenges and Opportunities

Nandini Sen
https://orcid.org/0000-0002-2844-6456
Heriot Watt University, UK

ABSTRACT

This chapter aims to create new knowledge regarding artificial intelligence (AI) ethics and relevant subjects while reviewing ethical relationship between human beings and AI/robotics and linking between the moral fabric or the ethical issues of AI as used in fictions and films. It carefully analyses how a human being will love robot and vice versa. Here, fictions and films are not just about technology but about their feelings and the nature of bonding between AIs and the human race. Ordinary human beings distrust and then start to like AIs. However, if the AI becomes a rogue as seen in many fictions and films, then the AI is taken down to avoid the destruction of the human beings. Scientists like Turing are champions of robot/AI's feelings. Fictional and movie AIs are developed to keenly watch and comprehend humans. These actions are so close to empathy they amount to consciousness and emotional quotient.

INTRODUCTION

Following Turing, A.M., (1950, p-433) I could have proposed the question, "Can machines think?" Turing's famous paper (1950) explains the meaning of the terms, "machine" and "think." I digress from his proposal a bit and further ask if machines can think ethically or if they can dream. The "machine" of the master mathematician has now been transformed into artificial intelligence and much progress has been done in this field. Now the question is that, how far the experts and scientists of artificial intelligence could add the intriguing value of the ethics in their experiments and brought the field of AI in the purview of gender, race, and ethics for the benefit of the human being and the animal kingdom. In

DOI: 10.4018/978-1-7998-6985-6.ch028

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the Imitation Game Turing (1950) was trying to observe what were be the changes which could happen if a machine intervened in the process. He is further asking if the investigation is crucial for answering the question – "Can machines think?"

Even before Stephen Hawking and Elon Musk started to structuralise the world of artificial intelligence in 1942, the science fiction writer Isaac Asimov wrote the famous "The Three Laws of Robotics: A moral code to keep our machines in check. And the three laws of robotics are: a robot may not injure a human being, or through inaction allow a human being to come to harm. The second law, a robot must obey orders given by human beings, except where such orders would conflict with the first law. And the third, a robot must protect its own existence as long as such protection does not conflict with the first and the second law." Apparently these three laws are sufficient for constructing the basis to work from to develop a moral fabric inside the AI. (Asimov, 1990; A Britannica Publishing).

After WWII, several people started to work on intelligent machines. The British pioneer in this field Alan Turing delivered a lecture on AI in 1947. He also is the first to decide that AI was best studied by programming computers rather than by building machines. By the late 1950s, there were many explorers on AI, and most of them were developing their work on programming computers. Alan Turing's 1950 article Computing Machinery and Intelligence discussed conditions for considering a machine to be intelligent. He argued that if the machine could successfully pretend to be human to an expert observer then it certainly should be considered intelligent. This test would satisfy most people but not all philosophers. The observer could interact with the machine and a human, and the human would try to persuade the observer that it was human, and the machine would try to fool the observer (Anderson, 2017; McCarthy, 1990).

The philosopher John Searle said that the idea of a non-biological machine being intelligent is not rational. The philosopher Hubert Dreyfus said that AI is a kind of utopia. The computer scientist Joseph Weizenbaum said that the idea of AI was dirty, anti-human and unethical. Various people have said that since artificial intelligence hasn't reached human level by now, it should not be imagined. The highly evolved field of robotics is producing a huge range of machines, from autonomous vacuum cleaners to military drones to entire factory productions. At the same time, artificial intelligence and machine learning are lagging behind the software that helps us daily, whether we're searching the internet or being given government tasks. In addition, many people say that AI is not a good idea. They even suggest remembering of films like Terminator and the Matrix where AIs had mainly inflicted destruction and/or exploitation of the mankind (Trent. (2012). These developments are rapidly leading to a time when robots of all kinds will become prevalent in almost all aspects of society, and human-robot interactions will rise significantly (Anderson, 2017); (McCarthy, 1990).

This chapter based on the above arguments will examine a logical progression in the evolution of AI. It will present a set of balanced arguments including the debates on the anticipated positives and the apprehended evils. These perhaps exist in the concepts if AI is necessary in human civilisation. It will take us through few sections mainly: literature review; discussion about fictional and filmic AI characters; comparison of real-life AI characters and fictional/filmic AI; ethnography; analysis and conclusion including future research.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

LITERATURE REVIEW

Asimov's laws are still mentioned as a guiding principle of the growth of the concept of artificial intelligence (AI). The South Korean government even proposed a Robot Ethics Charter in 2007 reflecting the laws. But assuming how much AI has changed and will continue to develop in the coming days, the chapter asks how these systems could be updated for a 21st century form of AI (Anderson, 2017). Asimov also examined how a machine could easily solve a mathematical problem but rejected to write a poem. On the other hand, Turing rightly trusted that a machine could solve the mathematical problem whereas the human being could give away if he had to solve a difficult problem. The AI machine could solve a chess problem within seconds, however, declared that he would not be able to write a poetry (Turing, 1950). Turing (1950) had tried to grasp the complex issue if machines or AI could perform better or differently from human being while undertaking a problem. He had also discussed if a machine had free will. It was tough to discern if the machine/AI had a random element or they could choose their programmes depending on the digits. (Turing, 1950, pp-433-460). Turing strongly believed, even if it was debatable about the thinking capacity of machines in 1950, however, in the later period of the century general people will not be able to argue about the machines' reasoning power. Turing is subtly critical about human beings that they could not go beyond their own superior intellectual capability. He wanted to establish a machine's reasoning ability. He though extended his view that apart from human beings, others could not have a soul as allocated by the God. He was not clear whether a machine could have a soul or not. That became the concern of the moral fabric. Machine is created by man. If man puts a soul inside the machine, then doesn't man not taking the role of the God (Turing, 1950).

Turing following Professor Jefferson's Lister Oration for 1949 agreed that a machine, until could scribble a poem or compose music inspired by thoughts and feelings and not by chance that machines will not attend or equalize human brain-that is, not only write it but know that it had written it. No AI could realise happiness at its successes, grief when its internal connections are down, be warmed by appreciation, be made depressed by its mistakes, be overwhelmed by sex, be desperate when it cannot get what it desires. This is directly taking the research to the question of consciousness and the mystery behind the machine or the AI. Turing clearly points out that the AI scientists need to construct kind, resourceful, beautiful, friendly, innovative, humorous machines. The AI should be able tell right from wrong, make errors, fall in love, relish delicious food, making attempts to make someone fall in love with it, taking lessons from experience, use words properly, can think about itself, have as much heterogenous behaviour as a man. These are possibilities of the near future, rather than Utopian dreams.

Copeland (2004) pointed out that Turing was not the only theoretical mind to have produced an idea basic for the development of computers: he was also involved in their construction. "It has been said that computing machines can only carry out the processes that they are instructed to do. This is certainly true in the sense that if they do something other than what they were instructed then they have just made some mistake. Turing, along with his colleagues, cracked the Enigma cipher, shortening WWII by an estimated two years (Copeland, 2004). When Alan Turing was cracking intercepted coded messages from the Nazis during World War II, he probably could not imagine that his computational discoveries could be used for identifying changes in patterns of ecological data as well. Turing was trying to recognize regularities in the coded messages of Nazi Germans and come up with a ''program'' that would transform the messages to their decoded content (Copeland, 2004).

Huang and Rust (2018) describe four levels of intelligence that AI systems can exhibit. While AI is not a new technology, given its conceptualization in the 1940s (Copeland 2004), it is still considered to

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

be an emerging technology since the techniques used to implement it are radically evolving (Stahl, et al., 2021). Additionally, companies are currently investing significant resources in the development and proliferation of AI. A simple question that comes to mind when scientists tried to define this concept is the following: is it a concept that can be understood in terms of a mechanism (or an algorithm) that generates certain (human-like) behaviours or is it a matter of a human perceiving an agent as intelligent? Although Turing's original intention in 'Computing Machinery and Intelligence' was to explore whether a computer can 'imitate a brain' (Turing and Copeland 2004), he then admitted to be skeptical as to how the intelligence of a machine was to be perceived: 'The extent to which we regard something as behaving in an intelligent manner' he noted (Turing 1950) 'is determined as much by our own state of mind and training as by the properties of the object under consideration'. In other words, Turing suggests that 'intelligence' relates to how we conceptualise it.

There have been a few notable scientists working for research centers and universities who have perceived the central concepts that have guided to technological initiatives in both the UK and the USA. Among these standouts is Alan Turing, the inventor of the Turing Machine, which set down the mathematical basics for the project of modern digital computers. The cryptography work of Turing in the Enigma project at Bletchley Park near London was crucial for shortening the duration of the Second World War. (Copeland, 2004).

Public worry will be a huge problem for AI to overcome. Understanding society's long-term tension of self-aware automatons should be a consideration within AI laboratories, especially those specializing in fully autonomous humanoids. Asimov anticipated this public terror and proposed the Three Laws of Robotics to reduce the fear. Three Laws are not practical from a general sense even though the ethical issues involved are crucial in the scientists' minds. As the contemporary idea of AI was still being refined, Asimov envisioned a day when humanity would be served by a host of AI. However, he knew that apprehension would be the greatest hurdle to success and, consequently, implanted all of his fictional robots with the Three Laws of Robotics. Above all, these laws served to protect humans from almost any abstracted risks. Asimov believed that humans would put safeguards into any potentially hazardous instruments and visualised AI as just advanced machines. Asimov considered that his Three Laws were more than just a literary plan; he could realise scientists engaged in Artificial Intelligence (AI) research had taken his Laws to heart, (Asimov1990; Britannica). Knowledge of the Three Laws of Robotics appears to be homogenous among AI researchers, there is the encompassing attitude that the Laws are not implementable in any meaningful sense. With the field of Artificial Intelligence now 70 years old and the extensive use of AI products, it is time to rethink about Asimov's Three Laws in terms of applications to implementation and calculating the underlying fear of unmanageable AI (Copeland, 2004).

CONTEXT OF THE FICTIONAL ROBOTIC AND AI CHARACTERS

Many storytellers and scientists have used similar rational to establish that only humans possess a soul. To interfere in this area is to intrude in God's own territory. This anxiety of man stopping, through technology, into God's zone and being unable to control his own creations is referred to as the "Frankenstein Complex" by Isaac Asimov in several of his essays (most notably (Asimov 1991). The "Frankenstein Complex" (Mc Cauley, 2007) is alive in Hollywood films rekindling the idea of love/hate through a line of productions that would never be forgotten: Terminator (all three); I, Robot; A.I.: Artificial Intelligence; 2001: A Space Odyssey; Blade Runner; Short Circuit; Electric Dreams; the Battlestar Galactica series;

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Robocop; Metropolis. These films originate from Sci-Fi literature. The important theme involves is that AIs attempt to harm people or even all of humanity. The current generation has seen a sudden and fantastic flowing in philosophy and science of the mind; a great intellectual revolution is happening in terms of AI. The central drama is AI, the fresh attempt to make computers think. In 1980's and 1990's people doubted the seriousness of a thinking machine. On the other hand, many scientists thought like Turing (1950) that it's only a matter of time, computers with minds will be possible, if interplanetary travel is endeavored. Scientists believed that AI psychology, philosophy, and linguistics, are now possible and can be established as super specialties. There is one fear not to recognise AI as a sensitive mind because of our anthropocentricism or human chauvinism is ingrained in our intelligence (Haugland, 1989).

Doug Blank (2006), Associate Professor of Computer Science at Bryn Mawr College, commented in a similar vein that trouble is that robots don't have clear-cut symbols and rules like those that must be imagined necessary in the sci-fi world. Most robots don't have the ability to look at a person and see them as a person (a 'human'). And that is the easiest concept needed to follow the rules. In the second novel of Asimov's Robots Series, The Naked Sun, the main character, Elijah Baley concludes that a robot could inadvertently disobey any of the Three Laws if it is not aware of the full consequences of its actions (Asimov 1957). While the character in the novel rightly conclude that it is impossible for an AI to know the full consequences of its actions, there is never an exploration of exactly how hard this task can be. (Blank 2006). Aaron Sloman, Professor of Artificial Intelligence and Cognitive Science at The University of Birmingham, described the issue in a way that gets at the sheer vastness of the problem: Another obstacle involves potential contradictions as the old utilitarian philosophers found centuries ago: harming and benefits by AI is relative in nature (Sloman 2006; Sloman 2006). This relativity in terms of harming and benefit which can be performed by an AI is complex (McCauley, 2007).

Isaac Asimov (1991) coined the term "Frankenstein Complex" to describe the fear that the public has towards humanmade technologies when they invade the realm commonly considered to be God's domain. This complex demonstrates some of the historical evidence of this fear and provides reasons why it is unfounded. Finally, it suggests a way for the AI community to help in reducing the public's fear of a prowling species of aggressive robots. Hollywood and the media worsen the issue while some well-known authors and scientists lend reliability to it. This apprehension is much older than the current concepts of AI and robotics. Asimov recognized this deep misinformation of technology and created the Three Laws of Robotics intended to project how these complex machines could be made harmless. For AIs to be accommodated in everyday society, scientists must examine the underlying apprehension of intelligent AI and reevaluate Asimov's response (McCauley, 2007). Not because he understood how to build robots, but to discover if there are lessons to be learned from his insights into society's response to intelligent technology. Asimov knew that fear would be the greatest barrier to success and, consequently, implanted all of his fictional robots with the Three Laws of Robotics. Above all, these laws served to protect humans from almost any perceivable hazard. Asimov believed that humans would put safeguards into any potentially dangerous tool and saw robots as just advanced tools. The word "robot" comes from the Czech word "robota" meaning 'drudgery' or 'servitude' (Jerz, 2002).

In much of science fiction since early times, the story is about artificially created workers that ultimately rise to overthrow their human creators. Even before Capek's (McCauley, L., (2007)) use of the term 'robot', however, the notion that science could produce something that it could not control had been explored most intensely by Mary Shelly under the guise of Frankenstein's monster (Shelley, 1818).

The full title of Shelley's novel is "Frankenstein, or The Modern Prometheus." In Greek mythology Prometheus brought fire (technology) to humanity and was, consequently, soundly punished by Zeus.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Where AIs are concerned, the fantasies that people can most readily recognise with, those that capture their imaginations and explore into their thickest fears, involve the displacing of humanity by its metallic offspring (Stephen Horne, 2021). Even well-respected scientists in both academia and industry have expressed their belief that humans will engineer a new species of intelligent machines that will replace human race. Ray Kurzweil (1999; 2005), Warwick (2002) argue that robotics, genetic engineering, and nanotechnology pose a unique kind of threat that the world has never before faced, "robots, engineered organisms, and nanobots share a dangerous amplifying factor: They can self-replicate. A bomb is blown up only once – but one bot can become many, and quickly get out of control." Clearly, Joy is expressing the basis of why the public at large continues to be grasped by the Frankenstein Complex. (Joy, Bill Joy, co-founder of Sun Microsystems, expressed in a 2000 Wired Magazine article (Joy, 2000)

The probability of a robotic uprising that destroys or subjugates humanity is quite low (Singh Roy, 2021). The primary argument that robots will take over the world is that they will eventually be able to design and manufacture themselves in large numbers thereby activating the inevitability of evolution. Each robot produced must have all the knowledge, capability, resources, time, and motivation to build more robots, otherwise, the evolution goes nowhere. Humans could build a robot or AI with the sole task of designing and building a new AI that is better at designing and building AI which builds another AI, etc. (McCauley, 2007). Asimov (1991) imagined human beings as the apprentice of the God in creativity when he is constructing the AI and robots. He is skeptical about the human's smartness in using the machines. Human beings submitted their souls to the Gods for the purpose of safety. In contrast to Turing (1950), Asimov (1991) imagined that human beings taking over creative powers of God or borrowing them to build their own superiority over nature. At every step Asimov expressed his doubts if smart people will wisely use the machines. Even in 1991 Asimov felt a pessimism that AI/Technology had carried away civilisation in a negative way. He even was frustrated when he put up the following question, "Do we not see that our Technology has run away with us and is slowly and inexorably destroying the environment ant habitability of the planet?" (Asimov, 1991).

Asimov (1991) directly said that human beings could strategise the whole matter by creating and artificial human being. Unlike Turing (1950) Asimov challenged the problem of creating artificial human being in different ways, even suggesting that AI engineers could be punished. Here we can take up the example of the story of Frankenstein by Mary Shelley (1818) where the creature created by the scientist became a monster and tried to punish the creator. The Artificial creature got trapped inside a world of loneliness. He was abandoned both by the audience and the creator to his fate. Asimov (1991) clearly advised his readers that the creator should be punished instead of the human-like creature. Asimov (1991) also introduced his readers to the concept of slave related to the 'robot' following Karel Kapek (play RUR). It is true that the concept of AI was dreams of human beings which were reflected in many stories. Mythical robots were used by the Gods or human beings under the guidance of the Gods. Shelley (1818) exposed very well that the scientist was not smart or wise enough to deal with this artificial being. He unempathetically treated the creature and in retaliation the creature killed everyone in Frankenstein's family including Frankenstein. Maybe Shelley (1818) even compared the God with this scientist. God clearly failed to control human beings from doing criminal acts and other sins. Shelley (1818) examined in her novel how God had also abandoned the human beings to fend for themselves and had even posed a question what would happen if the artificial creature was filled with culture and intellect and could talk like human beings. She pointed out that the artificial creature became miserably sad. He is unable to control his mind. Ultimately, he became dangerous for the human civilisation. The artificial being instead of finishing himself escaped to the Arctic. Asimov could envisage the future of

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the collaboration of the AI and human beings. Though scientists can confront the whole concept of AI, robots, or humanoid, at the same time they can develop an aura of romanticism. In fact, there are some things humanity is not meant to know. Human beings could find a solution to this unfortunate position of the AI. In future there would be lot of opportunities to develop an optimism through the wiser and balanced philosophy of handling AI ensuring safety of the broader humanity and human beings.

DISCUSSION AROUND AI CHARACTERS/ ROBOTS IN FICTIONS AND MOVIES

Fictional and movie AI characters can take the human beings through a mysterious adventure, sermon, or vision. Human beings, perhaps, can be overwhelmed by the intense abstraction and their strong narratives of absurdity, well- articulated though unresolved speculations on the origins and destinies of human life. They can as well grip us inside an engagement projecting their visual starkness, serenity, wonder, and euphoria at both the earth and the heaven, maybe beyond the infinite (Geraci, 2007). In science-fiction literature and film, human beings are simultaneously apprehensive and allured in the presence of intelligent machines. This is an experience that estimates the mystical experience science fiction offers. It is an empirical support for Anne Forest's (Forest, 1998) claim that human beings feel fear and fascination in the presence of advanced robots. The human reaction to intelligent machines shows that human beings in many respects have elevated those machines to divine status. Twentieth-century Euro-American cinema contradict our trust in the unambiguous potential of technology. *Metropolis*, *Forbidden Planet*, *Terminator*, and *The Matrix* show that human beings' relations with artificial intelligence complement their relations with the God. This signifies that robotic technology projects a particular kind of sanctity in the twentieth century. Technology promises salvation on one hand while creating fear with the other. This concurrence of opposites appears most prominently in depictions of intelligent robots (Geraci, 2007).

Lang in his film Metropolis (Geraci, 2007) shows an imagery in a scene where the protagonist, Freder, the son of the leader of Metropolis, wanders below ground while following a girl. In the underworld, he sees the machines go haywire after a worker collapses from fatigue. The mechanical failure results in numerous deaths, which the son sees in a vision as the sacrifice of children to the Babylonian god Moloch. In this powerful scene, Freder becomes aware of the dark side of a technology that has always benefited him. By superimposing the vision of Moloch over the machine, Lang shows how it has become a god, ruling over the subject workers who devote themselves utterly to its care. The allure and dread of technology continues in more contemporary films, from the *Terminator* series to the *Matrix* trilogy. In these movies, intelligent machines threaten the human species with extinction, real and virtual, but remain vital to human survival. Without Arnold Schwarzenegger's reprogrammed Terminator machine, the protagonist of *Terminator 2: Judgment Day* (1991) and *Terminator 3: The Rise of the Machines* (2003) would not have survived.

The research considers the foundational figure in science fiction to be Mary Shelley (1818), which would make Frankenstein's monster something of a foundational robot. In *Frankenstein: A Modern Prometheus* (Shelley, 1818), Frankenstein's monster, a living being made through scientific expertise, barely resembles the shiny metallic robots of many science-fiction authors (and especially films) but comes closer and closer to reality as scientists work to create artificial muscles, skin, and other tissue to replace the inefficient actuators and joints of contemporary robotics (Horne, 2021). Frankenstein's monster represents a hope for technological progress gone awry; the machine monster becomes a monster, indeed, and flees civilization (Brantinger, 1980). While *Frankenstein (1818)* marks a crucial moment

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

in literature, it wasn't until the twentieth century that true robots found their way into popular culture. Although Shelley's work tends toward the disastrous, twentieth-century science-fiction robots have largely been more complex (Horne, 2021).

*I, Robot (*Asimov, 1950*)* concludes with a story about the Machines—advanced robots that control Earth's economy for the maximum benefit of human beings. The Machines' immense power over Earthly life and their ability to fashion a harmonious Eden of it make them gods (Thomsen 1982). It transpires that the Machines are systematically eliminating human beings from positions of responsibility because those human beings mistrust the Machines and are therefore perceived as a threat to human well-being. The Machines' domination of human life means the reduction of humankind to mere instrumentality but also the possibility of real human happiness. Damnation and salvation are intertwined just like fear and fascination. Asimov ultimately believes that human beings will retain a control over their creations (Klass 1983), but he finds the wholehearted endorsement implicit in the *Star Wars* treatment of R2-D2 impossible to maintain.

The perfect 21st-century female looks like a smart model though a bit odd with the open -ended look. In "Ex Machina," Alex Garland's clever futuristic film about old and new desires, the female in question is a robot called Ava, a name suggestive of both Adam and Eve. Ava has a serene humanoid face and the expressive hands and feet of a dancer, but also the transparent figure of a visible woman anatomy model. Beautiful and smart, sleek, and stacked, Ava is at once decidedly unsettling and safely under lock and key, which makes her an ideal posthuman female (Manohlia, 2015). This prepares Ex-Machina a creative story, except instead of God repurposing a rib, the story here involves a Supreme Being who has built an A.I., using a fortune he's made from a search engine called Blue Book." The trouble with thinking machines, Wittgenstein writes, isn't that we don't know yet if they can do the job, but 'a machine thinks (perceives, dreams)' seems somehow nonsensical." And it seems so because such a machine is not (yet) known to us. Ava looks at once familiar and new, distinctly human, and thoroughly machined, only compared with the robot in "Metropolis". Ava excels the human-AI divide, she defies categorization because of the radical autonomy she shares with the queer sisters lived by Scarlett Johansson in "Her," and "Under the Skin". (Stephen Horne, 2021). These are the new heroines: totally hot, bracingly cold, powerfully sovereign, and posthuman. (Dargis, Manohlia., (2015).,

It's just a lump of iron. Why this attachment? A question asked of taxi driver Bimal (Kali Banerjee), the central character of Ajantrik, regarding his dedication to his very old and battered 1920 taxi, called Jagaddal. This taxi driver's relationship to his car somehow draws a parallel to Ghatak's own attempts to explain what it is about the cinema that draws his commitment. Ajantrik is a movie where a mingling of the human and the mechanical elements take place and is portrayed superbly when Bimal pauses pensively to say: "That I'm a machine. I like the smell of burnt gasoline. It makes me high…" The connect Bimal feels towards his taxi announces a profoundly human attachment and dedication motivating him. Bimal holds onto his car, Jagaddal, for fifteen years, against the prevailing trend among his peers who gave up old cars and upgraded to new ones. Bimal calls those new cars, 'fashionable whores.' Jagaddal thus is elevated to the status of a human being and Ritwik Ghatak brings out this human element through his camera, sights, and sounds, epitomised by his attention to frequent autonomous movements of Jagaddal's headlights. Sounds of drinking and exhalations of satisfaction exude from the car among descriptions of Jagaddal's health and durability. According to Bimal, in comparison with other cars, Jagaddal never 'catches cold' or 'gets tummy aches.' That Bimal believes in Jagaddal's independent agency is summarised in the final test of the car's strength, after it has received new parts (Information Desk, 2021). 'I've pampered you enough,' Bimal warns, dropping several large boulders and as the car collapses or

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

'dies,' Bimal smashes the windscreen and bursts into tears, his head resting on the steering wheel. And thus, Ritwik Ghatak superbly excels beyond the 'man and the machine' (Information Desk, 2021)

Meditations on the philosophical and social conundrums posed by Asimov's most celebrated invention: The Three Laws of Robotics. Asimov's stories have become such an established part of robot folklore that his three laws have shaped some of the fundamental thinking in the development of real-life robotic engineering and artificial intelligence. In the movie, however, much of the action is motivated by robots that have in some way been programmed to evade these laws, under which circumstances their chests conveniently start glowing red. Asimov's aim with his robot stories was to explore the consequences, and in particular the loopholes, of the rational scheme that he had established (Bell, 2004).

Transhumanists believe that human race will soon undergo a technological evolution into a new and superior form. While there is no single template for transhumanism's imagined future, there are several recurring motifs, such as enhanced cognition, improved bodies, and extended lifespans. Sometimes the emphasis is on enhanced biology, sometimes on the supplementation or replacement of the body by technology. the futures imagined in science fiction sometimes come true. Science fiction writers told stories about going to the moon before anyone went there. They told stories about artificially intelligent machines before these were invented. And so on. This means readers may try to rummage through science fiction for visionary images, especially of future technologies. The dystopian vision of Altered Carbon might, for instance, be construed as a warning that we need "to be thinking about the cost of pursuing technological immortality" (Miller, 2018). Westworld, a serial, challenges scientists to consider the difference between being human and being a AI. From the beginning of this new serialisation on HBO scenes of graphic human-on-robot violence became popular. But the robots in Westworld have more than just human-like physical bodies, they display emotion including extreme pain, they see and recognise each other's suffering, they bleed and even die. What makes this acceptable, at least within Westworld's narrative, is that they are just extremely life-like human simulations. The disturbing message, echoing that of previous sci-fi classics such as Blade Runner and AI, is that machines could one day touch human ability as to be indistinguishable from human beings – not just in intellect and appearance, but also in moral terms (Prescott, 2016).

But there are relatively few about how they might cooperate with humanity (Asimov got in here early on this and remains one of the few). The hypocrisy is that this trend suggests that it's fine for governments to monitor its citizens and for corporations to analyse social media feeds (even using software bots), but an AI shouldn't. It's like saying that you're happy being screwed over, but only by a political system or another mammal, not a computer. One solution, therefore, is to consider how to limit AIs and teach them human ethics. But if we "train" AIs to have ethical behaviours, who do we trust to train them? To whose ethical standards? Public perceptions of AI will be governed by just this sort of mistrust and suspicion, fostered by such public debacles and by the broadly negative view evident in much science fiction. But what such examples perhaps reveal is that the problem with AI is not that it is "artificial", nor that it is immoral, nor even in its economic or social impact. Perhaps the problem is us (Slocombe, 2016). Despite the "long time ago," Star Wars is picturised in the future in terms of how we conceive of it. The human society narrated is one that features more advanced technology than we currently have, a technology that reflects our trajectory in transportation, telecommunication, military technology, artificial intelligence, and robotics. So, what kind of future does Star Wars show us? Well, a familiar one, to put it mildly — one that despite the spaceships and laser-swords is more reflective than speculative. "From a social point of view most SF has been incredibly regressive and unimaginative.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Instead of comparing it to the works of H.G. Wells or Isaac Asimov or Kim Stanley Robinson, we can compare Star Wars to the works of George MacDonald and J.R.R. Tolkien and Neil Gaiman, and there we find a much more favourable lens by which to understand and appreciate what the Star Wars universe accomplishes. This isn't SF – it's fantasy in outer space, and, more than that, it's good fantasy (Deman, 2017).

Asimov's stories as allegories of slavery point out that the biologist of the robot would be dealing with himself, and not only because he is the only instantiation but also because that's what thinking humans do. Asimov explicitly refers to Andrew's first scholarly achievement, namely a history of robots: he wants to explain how robots feel about what has happened since the first ones could work and live on Earth. Only a robot can tell what it is like to be a robot and telling this realizes the history of robots from an inside perspective and thus again a narrative of the life of someone who is supposed not to be living. Robots do not have a history, they live a timeless life, but if robots can produce narratives of their lives, they prove to be human (Blum, 2016). In "The Bicentennial Man" (Asimov, 1976), Asimov rejected his own Three Laws as a proper basis for Machine Ethics. He believed that a robot with the characteristics possessed by Andrew, the robot hero of the story, should not be required to be a slave to human beings as the Three Laws dictate. He, further, provided an explanation for why humans feel the need to treat intelligent robots as slaves, an explanation that shows a weakness in human beings that makes it difficult for them to be ethical icons. Because of this weakness, it seems likely that machines like Andrew could be more ethical than most human beings. "The Bicentennial Man" gives us hope that, not only can intelligent machines be taught to behave in an ethical fashion, but they might be able lead human beings to behave more ethically as well (Anderson, 2007).

Robots have become as commonplace as vacuum cleaners, gene-editing is the norm and biotechnological advances are close to recreating unique human beings. Amazon recommends is just the beginning. "In the era of big data, we might start to be able to rebuild somebody's character so that after they've died, they can still carry on, figuring out what they'd order next online, which concert they'd like to go to and what they would have said at the breakfast table if you had read them the latest headlines," Ishiguro continues (Allardice, 2021). He deliberately didn't read either the recent Ian McEwan novel *Machines Like Me* or Jeanette Winterson's *Frankissstein*, which also take on artificial intelligence. From very different angles. Klara (Ishiguro, 2021) is a sort of robotic parent, "Terminator-like in her determination to look after Josie", but she is also a potential surrogate child: when Josie gets sick, Klara is being programmed to take her place. "What happens to things like love in an age when we are changing our views about the human individual and the individual's uniqueness?" Ishiguro asks. "There was this question – it always sounds very pompous – about the human soul: do we actually have one or not?" (Ishiguro, 2021). For the first time, Ishiguro is beginning to fear for the future, not just the consequences of climate change, but other issues raised in *Klara*: artificial intelligence, gene-editing, big data – "sorry to bang on about this" – and their implications for equality and democracy. "The nature of capitalism itself is changing its model," Ishiguro says. "I do worry that we are not in control any more of these things." (Allardice, 2021).

The film script, written by Kubrick and Arthur C. Clarke whose books and stories the movie was based on, and directed by Kubrick, is a multisensory ode to cosmic mystery, fate and the future. Kubrick's *2001*, like most science fiction, problematizes humanity's relation with technology. And, like most science fiction, it is not so much about the future, whatever that could mean, as the future as seen from the present of the filmmaker's eyes and thoughts. The cold calculating nature of the characters is the audience's coldness; their lack of surprise at their surroundings is the lack of surprise at the audience. The film's characters are lost (Pope, 2003). The film's characters are lost and have surrendered to their

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

crypts in an all-pervasive melancholy. Narratives of transcendence and/or Apocalypse spring. HAL, the AI, goes mad. Is Kubrick suggesting we regain control of the AI, or that human beings renegotiate their relationship with technology, accepting their contamination by it, or their co-dependence with it? Given the nature of the ending, there can be no conclusions here, though it should be noted that even after Bowman kills HAL, he still employs the AI, goes for a ride with it -- beyond the infinite (Overbye, 2018).

Consumers of devices such as Apple Computer's Siri or Amazon's Alexa are acquainted with the way AIs communicate with humans with voice- and thought recognition. The clarity of diction and thought compositions of these machines remind us of the exploits of the supercomputer Hal 9000 of Arthur C. Clark's 2001: A Space Odyssey, as depicted in Stanley Kubrick's movie of the same name. Hal's voice quality was certainly not as human-like as Alexa's, for instance, but Hal demonstrated distinct emotive qualities not yet seen in these familiar commercial devices (Swagatam Sen, 2021). We can argue that artificial intelligence of digital technology thus far is superfast in information processing and all the exploits of the extant machines merely are a response to information fed as instruction (Swagatam Sen, 2021). In this sense, interaction with Alexa is not a conversation, as Alexa cannot speak about anything what she has not been supplied as data or has not accumulated by listening to her human consumers (Sen, S., and Singh Roy, 2021). On the other hand, when we order an article of commerce through a vendor's website, frequently you receive suggestions with this kind of message: buyers who ordered this are also interested in the following. A clever marketing trick alright, but behind it is an algorithm that looks at your profile of preference and makes an informed suggestion. Consumers need to think of the consequence of using these awesome machines (Horne, 2021). The scariest consequence of artificial intelligence has been sounded by Elon Musk of Tesla Motors and supported by Stephen Hawking. Since technologies have social consequences, what would happen if we were able to achieve driverless public transportation, robots serving in the supermarkets and restaurants and similar developments throughout social activities? Or, just a few very technology-savvy individuals will control the rest of us? The answers to the questions raised here will slowly but surely be revealed in future (Sikdar, 2018).

The late Stanley Kubrick's film 2001: A Space Odyssey *portrayed a computer, HAL9000, that appeared to be a conscious entity, especially given that it seemed capable of some forms of emotional expression.* Looking back on how humans interact with a supercomputer in *2001: A Space Odyssey*, we are led to conclude that many forthcoming challenges of dealing with synthetic emotions in AI will be more social than technical. Although a work of fiction, this film can nevertheless alert us to specific pitfalls that can occur when human actors embark on emotion-laden interchanges with computers. Computers, too, may be prone to misinterpretation of human emotional cues. Although this is not unusual among human actors, the high-speed logical processing and problem solving of computers creates a far greater level of asymmetry in situations of emotional exchange (Nofz, et al, 2002).

ETHNOGRAPHY

I came across four people Tirtho Dasgupta, Swagatam Sen, Anusua Singh Roy, and Stephen Horne who are either Artificial Intelligence scientists or admirers of films and fictions based on artificial intelligence. I thought to interview and chat with them over phone on the subject in-depth to add clarity and qualitative dimensions to the research.

Tirtho Dasgupta, Film Maker and AI user-experience designer:

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

AI slowly entered the commercial world after the scientific world. I used Facebook very nominally as I had other things to do in my hometown. After coming to Scotland, I experienced loneliness and homesickness. These factors influenced me to go deeper into Facebook (FB). Rootlessness has provoked me to spend more time with FB. I dug FB for connecting myself with my childhood. Thus, FB AI gives opportunity me to reconnect with my nostalgia, roots, childhood, schools, and the friends related to these spaces. AI is giving me scope to connect with the common friend- school friends, local friends, feel-good cultural groups- to develop positive feelings, new friends, and to interact with new characters. I argue, exchange narratives, and share my feelings with new and older friends. At the same time political and religious sectors in FB provoked me to become emotionally agitative. I am addictive to FB and Twitter and cannot come out these unnecessarily evocative AI platforms. A deeply involved person like me is forgetting to empathise or listening to other people. However, I have now recovered and have come under control. I am questioning that if the AI spaces are safe or whether they are destroying our personal and friendly relationships. I observe that my younger friends are trying to experience instant gratification through these social media controlled by sensitive and super intelligent AIs. The business houses are controlling our souls, and moral fabric by simply playing algorithms. These AIs, through social business medias, are manipulating with our choices, even in terms of our sadnesses and happy moments. When someone close is dying we are sharing this intimate experience within an unsafe, very public space like FB. Since 2015-2014- the smartphone revolution created this selfie-obsession and breaking the Monalisa myth by making it only the background. We are becoming crazy to take a photo in front of that small painting and instantly sharing the photo with the public. I am getting out-of-proportionately bigger than that legendary photo of Monalisa. I am not appreciating the place or the creation- photo is not etched in my mind, mind is replaced by the prompt gratification. Social media is worrying me I wonder what kind of world we are creating for the future generation. AI hasn't matured enough to create its own subjectivity or subjective perspective- subjective decision. This is an enigma if AI can solve the acute problem of ethics and will be able to produce a matured, empathetic AI who will help and support to sustain humanity in this world.

Swagatam Sen: HSBC, Senior Manager, Research and Development, Compliance Analytics

I will emphasise on data witchcraft or the AI which can be used in medical science medicine profession. *2001: Space Odyssey*, it's a science fiction- one of the pioneers about AI. Hall is an elevated machine. It has a philosophical tone. He is asking what happens if AI has the consciousness. Hal comprehended that there are errors in the whole operation inside the spaceship. Hal is almost adopting human consciousness and emotions. At first Hal is trying to convince Dave to run the operation. Dave is not understanding Hal's logic. Hal, as if developing human instinct, can see through the human error and to make the mission successful is killing the scientist and others. Sen (2021), believes that scientists and consumers must trust AIs, create protocols, and can discover the optimal solutions to mitigate ethical issues.

Anusua Singh Roy, Statistician, Queen Margaret University, Edinburgh, UK

Singh Roy talks about the ethical issues related to gender, class, and community perspectives in the kingdom of AI. She is challenging how far Asimov's three laws of robotics are applicable in the modern world taking consideration of profitability and expanse of industrial territories although we sometimes forget to keep in mind about climate change. The considerate admirer of AI has concerns whether industries and science laboratories comprehend robots and AI as slaves of human beings or whether the AI/ robots will control human being in future. She said still we are clueless about the conflict of interests between AIs and human race. Following Turing's arguments (1950) Singh Roy genuinely wishes to know if AI/ Robots dream, watch nightmares, if can they participate in community activities or feel

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

any urge to be in kinship. She is curious to know if the AIs can reproduce their next generation without human intervention. She also stated that still now scientists are in darkness about the real quantification of the AI emotional capital. This lack of knowledge about AI emotions and consciousness may bring the extinction of human beings. AIs can reach a such a stage where they will not be able to control their rage and can kill human beings. In Marvel-movies- AI- Avenger's Age of Ultron- Ultron was initially peacekeeping programme- controlling violence but ultimately became uncontrolled. He gathered so much information that Ultron realised that to maintain peace human beings should be destroyed. He learned from surroundings that human beings are the source of all problem and sins. Scientists must rethink even if the human beings are erroneous and sinful that doesn't mean that they need to be made eroded from the face of the earth. Human race should be controlled otherwise they can destroy other human beings, or the humans must extinct. Sing Roy analyses the situation in a very humane way. It goes in line with how much advanced the robots and the AIs will be, this must be decided otherwise AIs will destroy human beings. The AI admirer suggests that robots and AIs can be filled in with ethical values so that AIs can differentiate good and bad or right and wrongs. AIs without humane qualities will remain as raw intelligence and AIs can become rogues. Scientists then will need to dismantle roguish AIs. Instead of taking down such expensive intelligent helpers, scientists can carefully craft the AIs with ethical and humane values.

HOW FAR CAN AI BE CREATIVE AND USED FOR PROFIT IN THE MAINSTREAM SOCIETY OR FOR INDUSTRIES

In 1980's the computers are excelling their programmers. Even the programme designers are always not able to predict how the AI would work after getting the inputs from the scientist. Chess machines are outperforming their programmers bringing up some brilliant moves that the engineers will never have imagined in their dreams (Sen, 2021). However, this dramatic invention is only a rearrangement of previously available matters. Thus, many AI experts (Sen, 2021) will reflect the whole process of AI as truly creative, free, or artistic. But at the same time other AI scientists will push to perform harder and more complex research to reproduce the true AI with an intellect which can create jealousy in the scientist's mind. It needs to be cultivated if current generation of AIs can feel the alienation, loneliness, or the melancholia of their legendary forefather Frankenstein (Sen, S. 2021).

Intense existential threats (Davies, 2016) from AI comes from misplaced attention on the possibility that such technology could develop consciousness. Recent headlines suggest that respected thinkers such as Bill Gates and Stephen Hawking are concerned about machines becoming self-aware. At some point, a piece of software will 'wake up', prioritize its desires above human beings and intimidating humanity's existence. We must realize that stopping an AI from developing consciousness is not the same as stopping it from developing the capacity to cause terror. It might be that consciousness, or our perception of it, would naturally come with superintelligence. That is, the way we would judge something as conscious or not would be based on our interactions with it. A superintelligent AI would be able to talk to the consumers, create faces that react with emotional expressions just like somebody we will talk to on Skype, or what's app or zoom. Maybe with conscious AIs would care about humanity more than unconscious ones would (Sen, 2021). In his book *Superintelligence* (Oxford University Press, 2011), the Oxford researcher Nick Bostrom (2011) describes many examples of how an AI could be dangerous. One is an AI whose main ambition is to create more AIs. With advanced intelligence and no other values, it

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

might proceed to seek control of world resources in pursuit of this goal, and humanity can be damaged. Another scenario is an AI asked to calculate the infinite digits of pi that uses up all of Earth's matter as computing resources. Perhaps an AI built with more laudable goals, such as decreasing suffering, would try to eliminate humanity for the good of the rest of life on Earth. These hypothetical processes are dangerous not because they are conscious, but because they are built without refined and multifaceted ethics. Instead of getting fanatical about consciousness in AI, scientists should put more effort into programming goals, values, and ethical codes. A global race is under way to develop AI. And there is a chance that the first superintelligent AI with ethical values will be the only one, scientists ever make. So that it can improve itself and start changing the world according to its own values and maintaining safety of the human kingdom

Turing (1950) argued that it must be possible to build a thinking machine since it was possible to build imitations of any small part of a man. He made the distinction between producing accurate electrical models of nerves and replacing them computationally with the available technology of vacuum tube circuits, this follows directly from his paper (1950), and the assumption that the nervous system can be modeled as a computational system. For other parts of the body, he suggests that \television cameras, microphones, loudspeakers", etc., could be used to model the rest of the system. This would be a tremendous undertaking of course." Even so, Turing notes that the so constructed machine would still have no contact with food, sex, sport, and many other things of importance to the human being. Turing concludes that the best domains in which to explore the mechanization of thought are various games, in that they require little contact with the outside world. (Brooks, 1991).

One potential opportunity of AI is in updating intellectual property policy for the AI. The consumers could, for example, reduce patent lengths for the use of digital technologies to ensure that the gains from AI-based innovations are shared broadly. There is a strong case for granting developing countries access to patented technologies: most patents are produced in rich countries and generate most of their income from sales and licenses within rich countries, so innovators would not incur significant losses if developing countries could use their technology for free or with limited royalties before their patents expire in high-income countries (Korinek, Stiglitz, 2021).

ANTICIPATED AI ACTIONS

Aissatou, from a rural village in Guinea, West Africa, couldn't add the essayists' phone numbers to her phone so they won't be able to stay in touch with them. She said, "I can't, because I did not go to school." Lacking a formal education, Aissatou does not read or write in any European or mainstream language. Aissatou's lack of schooling should not keep her from accessing basic services on her phone. The problem, as we see it, is that Aissatou's new phone does not understand her local language. Computer and AI systems should adapt to the ways majority of the people use language. West Africans have spoken their languages for thousands of years, creating rich oral history traditions that have served communities by bringing alive ancestral stories and historical perspectives and passing down knowledge and morals. Computers could easily support this oral tradition. Speech technology, however, does not "speak" any of the 2,000 languages and dialects spoken by Africans. Apple's Siri, Google Assistant, and Amazon's Alexa collectively serve zero African languages. Maybe this can be fixed by propagating a more inclusive language training and input in these AIs (Doumbouya and et al, 2021). In fact, the benefits of mobile technology are not accessible to most of the_700 million illiterate people around the world. Because illit-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

eracy tends to correlate with lack of schooling and thus the inability to speak a common world language, speech technology is not available to those who need it. For them, speech recognition technology could help linking the gap between illiteracy and access to valuable information and services from agricultural information to medical care which can be passed on. Groups with power over technological goods and services tend to speak the same few languages, making it easy to insufficiently consider those with different backgrounds. Speakers of languages such as those widely spoken in West Africa or other poor countries are unfairly underrepresented in research of companies and universities that have historically built-up speech-recognition technics. It is unfortunate that technological systems can fail to provide the same quality of services for diverse consumers, treating many communities as if they do not exist.

Labor-saving advances in AI may undo the gains from globalization and pose new challenges for economic development. The last fifty years have spurred unprecedented economic growth for many developing countries – brought about by the fast adoption of technological advancements – and has led to vastly shared increases in prosperity, lifting many people out of poverty. But there is no guarantee that this process will continue in future. Their (Korinek, Stiglitz, 2021) new working paper on "Artificial Intelligence, Globalization, and Strategies for Economic Development," challenges the long-standing assumption that technological progress will necessarily continue to advance broadly shared prosperity in developing countries. They argue that new developments in (AI) may create the perfect storm to topple technological advancement from broadly shared increases in living standards. Dozens of developing countries have escaped from poverty in recent decades through a strategy of export-led growth, fueled by manufacturing goods produced with their abundant supply of cheap labor or their richness in natural resources (Korinek, Stiglitz, 2021). But AI poses a triple threat to this strategy. "First, AI automates labor, which strips developing countries of their comparative advantage in cheap labor. Second, AI produces more output with fewer natural resources. And for nations that rely on exporting their natural resources to pay for food and other essentials, AI advancement could have deadly consequences. Third, AI operates in a "winner-takes-all" market." (Korinek, Stiglitz, 2021). AI scientists need to address these crucial problems and bring out concrete solutions to make the world fairer for majority of the people.

Metaphors inspired by differences of individual intelligence between humans and AIs will patent new inventions, publish groundbreaking research papers, make money on the stock market, or lead political power blocks. Fast AIs will invent capabilities that futurists commonly predict for human civilizations a millennium in the future, like interstellar travel. (Bolstrom, N., Yudkowsky, E., 2011). Asimov's laws are based on functional morality, which assumes that robots have sufficient agency and cognition to make moral decisions. It is debatable if Asimov's laws distract from capturing the diversity of robotic missions and initiative. Understanding these diversities and complexities is critical for designing the "right" kind of AI. Ironically, Asimov's laws really are human-centric because most of the initiative for safety and efficacy lies in the robot as an autonomous agent. Finally, while perhaps not as entertaining as Asimov's laws, experts can hope the alternative laws of responsible robotics can better communicate to the public about the complex mix of opportunities and challenges of robots in today's world (Murphy and Woods, 2009).

CONCLUSION

If we are serious about developing advanced AI, this is a challenge that we must meet machines are to be placed in a position of being stronger, faster, more trusted, or smarter than humans, then the discipline

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of machine ethics must entrust itself to seeking human-superior (not just human-equivalent) niceness. Although current AIs offer us few ethical issues that are already present in the design of cars or power plants, the approach of AI algorithms toward more humanlike thought foreshadows speculated complications. Social roles may be filled by AI algorithms, implying new design requirements like transparency and foreseeability. (Bolstrom, N., Yudkowsky, E., (2011). Scientists and consumers of AI may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. However, we have discussed in the section of **Anticipated AI Actions,** about AIs only with English as a language of communication, will be wicked in nature as majority of the people in the world from the poorer countries do not understand English. It is needed to innovate several different languages which can be put inside the AI so that AIs become universal and more humanitarian in nature (Mc Carthy, 2004).

As AI systems (e.g., Robots, chatbots, and other intelligent agents) are moving from being perceived as a tool to being perceived as autonomous agents and team-mates, an important focus of research and development is comprehending the ethical impact of these systems. What does it mean for an AI system to decide? There should be the moral, societal, and legal consequences of their actions and decisions. It will be difficult to hold AI accountable for their actions and to expect AIs will stop their learning capacities and can be controlled by their creators. Contrary to the frightening images of a dystopic future in media and popular fiction, where AI systems dominate the world and is mostly concerned with warfare, AI is already changing the everyday lives mostly in ways that improve human health, safety, and productivity (Stone et al. 2016). This is the case in domain such as transportation; service robots; healthcare; education; public safety and security; and entertainment.

The researcher challenges the concept of a general welfare. The countries which are benefiting from the AI systems are the minority rich countries who have the capability to invest huge amounts of money to innovate and cultivate such AIs and robots in their everyday lives. To ensure that those dystopic futures do not become reality, these systems must be introduced in ways that build trust and understanding, and respect human and civil rights. The need for ethical considerations in the building of intelligent interactive systems is becoming one of the most significant areas of research in the last few years and has led to several plans both whatever their level of autonomy and social awareness and their ability to learn, Theories, methods, algorithms are needed to integrate societal, legal and moral values into technological developments in AI, at all stages of development - analysis, design, construction, deployment, and evaluation. These frameworks must deal both with the poorer countries and richer countries' autonomic reasoning of the machine about such issues that we consider having ethical impact. AI reasoning should be able to consider societal values, moral and ethical considerations; weigh the respective priorities of values held by different stakeholders in various multicultural contexts; explain its reasoning; and assure transparency. Responsible Artificial Intelligence is about human responsibility for the development of intelligent systems along fundamental human principles and values, to ensure human flourishing and wellbeing in a sustainable world for both poorer and richer countries (Dignum, 2018).

Ten years ago, if scientists mentioned the term "artificial intelligence" in a boardroom there's a good chance you would have been laughed at. For most people it would recall sentient, sci-fi machines such as 2001: A Space Odyssey's HAL or Star Trek's Data. Today it is one of the hottest buzzwords in business and industry. So how has this change come about? Well partly it is due to the Big Data revolution itself (Sen, and Dasgupta, 2021). From healthcare to self-driving cars to predicting the outcome of

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

legal cases are all done by AIs, no one is laughing now! Here too the current research will allege that though no one is laughing at the huge technological development in the rich countries who can afford AIs in their everyday lives powered by AI in their smartphones, assistants like Apple's Siri and Google's Google assistants, self-driving autonomous cars, but people in the poorer countries are unable even to feed themselves properly and are suffering in their digitally divided worlds. Even in countries like UK the latest pandemic has clearly polarized between the rich and the working- class people in terms of the digital and AI usage and consumption (Marr, 2017).

Real anxieties that development of AI which equals or surpasses human brain but has the capacity to work at far higher speeds, could have negative implications for the future of humanity have been voiced, and not just by apocalyptic sci-fi such as The Matrix or The Terminator, but respected scientists like Stephen Hawking. This will further divide the world. Even if AIs don't force human beings to leave the earth or turn human race into living batteries, a less dramatic but still nightmarish scenario is that automation of labour will lead to profound societal change perhaps for the better, or perhaps for the worse. This comprehensible concern has led to doubting by several tech titans including Google, IBM, Microsoft, Facebook, and Amazon, of the partnership in AI. We need to trust that these groups will research and advocate for ethical implementations of AI, and to set guidelines for future research and deployment of robots and AI. The question arises here that whether these big AI giants will bring in the whole spectrum of problems like racial, gender, equality, culture, societal, and poverty divides in the kingdom of AI. If these crucial and important problems are not sorted and mitigated with sustainable solutions the world of AI will be completely meaningless.

The artificial intelligence (AI) and robotics communities face an important ethical decision: whether to support or oppose the development of lethal autonomous weapons systems (Singh Roy, 2021, Russel, 2015). There are ethical issues involved when efforts are made to apply normative ethical principles on non-organic intelligence like AI or consciousness uploads (human brain being transformed into inorganic computers which can imitate the original brain personhood). When human consciousness is uploaded onto a fast computer, the principle of subjective rate of time, for example, could be incompatible with human ethics. Many dignitaries like Stephen Hawking, Steve Wozniak, and the figureheads of AI at Facebook and Google, Yann LeCun, and Demis Hassabis have signed a petition of warning of a ''military artificial intelligence arms race" and calling for a ban on dangerous weapons and military conflicts (Kadircan, 2019). The research, in conclusion, will like to know how AI can act ethically while making endorsements to humans and express their ethical judgements effectively. These are the current foci of ethical human-AI interaction research. In addition, AI engineers and experts need to engage more with the ethics and decision-making scientist communities. Since such AI technologies as autonomous vehicles, and autonomous weapons, and cryptocurrencies are becoming a reality and affecting societies, a global and unified AI regulatory framework needs to be established as soon as possible to address the ethical issues by drawing on interdisciplinary expertise (Goldsmith and Burton, 2017).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Allan Turing: Life and Legacy of a Great Thinker. (n.d.). https://link.springer.com/chapter/10.1007/978-3-662-05642-4_13

Allardice, L. (2021). *Interview: Kazin Ishiguro: AI, gene-editing, big data… I worry we are not in control of these things anymore*. The Guardian. https://www.theguardian.com/books/2021/feb/20/kazuo-ishiguro-klara-and-the-sun-interview?fbclid=IwAR0ExKc5Jda-6TEWM-l-zJzBGYmFgpw0N-DWVrlzPfIl6Ehw3Dh1WPIaRMI

Anderson, M. R. (2017). *After 75 Years, Isaac Asimov's Three Laws of Robotics Need Updating*. The Conversation.

Anderson, S. L. (2007). *Asimov's "Three Laws of Robotics" and Machine Metaethics. Artificial Intelligence and Society, 22*.

Asimov, I. (1950). *I, Robot*. Del Rey.

Asimov, I. (1957). *The Naked Sun*. Doubleday.

Asimov, I. (1990). The Laws of Robotics. In Robot Visions. Academic Press.

Asimov, I. (1991). *The Lord's Apprentice. In The Ultimate Frankenstein*. Manhanset House.

Asimov, I. (n.d.). *Learn about Asimov's Three Laws of Robotics*. Britannica Publishing.

Bell, P. (2004). Monster or Machine? *Nature*.

Blank, D. (2006). *Robotics and Asimov's Three Laws. In The Frankenstein Complex and Asimov's Three Laws*. Association of the Advancement of Artificial Intelligence.

Blum, P. R., (2016). Robots, Slaves, and the Paradox of the Human Condition. In *Isaac Asimov's Robot Stories*. Czech Science Foundation at the project Between Renaissance and Baroque Philosophy and Knowledge of the Czech Lands within the Wider European Context.

Bolstrom, N., & Yudkowsky, E. (2011). *The Ethics of Artificial Intelligence*. Cambridge University Press.

Brantinger, P. (1980). The Gothic Origins of Science Fiction. *NOVEL: A Forum on Fiction, 14*(1), 30–43.

Brooks, R. A. (1991). *Intelligence without Reason*. Massachusetts Institute of Technology Artificial Intelligence Laboratory.

Copeland, B. J., & Produfoot, D. (2004). *The Computer, Artificial Intelligence, and the Turing Test. In Alan Turing: Life and Legacy of a Great Thinker*. Springer.

Dargis, M. (2015). Review: In 'Ex Machina', a Mogul Fashions the Droid of his dreams. *The New York Times*. https://www.nytimes.com/2015/04/10/movies/review-in-ex-machina-a-mogul-fashions-the-droid-of-his-dreams.html

Davies, J., (2016). Program good ethics into artificial intelligence. *Nature*, *538*, 291.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Deman, J. (2017). Star Wars is colonial fantasy: How our future imaginings are limited by our past. *The Conversation*.

Dignum, V., (2018). Ethics in artificial Intelligence: Introduction to the Special Issue. *Nature, Ethics and Information Technology*, *20*, 1-3.

Doumbouya, M., Einstein, L., & Piech, C. (2020). Why AI Needs to be able to understand all the World's languages: The benefits of mobile technology are not accessible to most of the world's 700 million illiterate people. *Scientific American*. https://www.scientificamerican.com/article/why-ai-needs-to-be-able-to-understand-all-the-worlds-languages/?fbclid=IwAR3sA7mY-yDukaLdu0dC9RKCzZMEdtubdG5WoGjmk8hc4pMSY6Fdzkbt7No

Forest, A. (1998). Cog, a Humanoid Robot, and The Question of The Image of God. *Zygon: Journal of Religion and Science*, *33*, 91-111.

Geraci, R. M. (2007). Robots and the Sacred in Science Fiction: Theological Implications of AI. *Zygon*, *42*, No-4.

Goldsmith, J., & Burton, E. (2017). Why Teaching Artificial Intelligence Practitioners Is Important. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), 4836-4840.

Haugland, J. (1989). *Artificial Intelligence: The Very Idea*. The Massachusetts Institute of Technology.

Huang, M., Rust, R.T., (2018). Artificial Intelligence in Service. *Journal of Service Research*, *2*(2), 155-172.

Information Desk. (2021). *Ritwik Ghatak's War between 'Man and Machine*. Get Bengal, Thinking Positive.

Ishiguro, K. (2021). *Klara and the Sun*. Faber and Faber.

Jerz, D. G. (2002). R.U.R. (Rossum's Universal Robots). Dover.

Joy, B. (2000). *Why the future doesn't need us*. Wired.

Kadircan, H. K. (2019). Medical Ethics: Consideration on Artificial Intelligence. *Journal of Clinical Neuroscience*.

Kevin, W. J. (2000). *Co-founder of Sun Microsystem-expressed in a 2000 Wired Magazine article*. Academic Press.

Klass, M. (1983). The Artificial Alien: Transformations of the Robot in Science Fiction. *Annals of the American Academy of Political and Social Science*, 171-79.

Korinek, A., & Stiglitz, J. (2021). *Artificial Intelligence Could Mean Large Increase in Prosperity- But only for a Privileged Few*. https://www.ineteconomics.org/perspectives/blog/artificial-intelligence-could-mean-technological-advancement-but-only-for-a-privileged-few?fbclid=IwAR0qxVBh0NYHEJLVJCoC3bynkESnLFgfn8QFZcIE63kGu06TIjMEbWDHadM

Kurzweil, R. (1999). *The Age of Spiritual Machines*. Viking Adult.

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking Books.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Manohlia, D. (2015). Review: 'Ex Machina', a Mogul Fashions the Droid of his Dreams. *The New York Times*.

Marr, B. (2017). The Complete Beginners' Guide to Artificial Intelligence. *Forbes*.

McCarthy, J. (1990). *What is AI? Basic Questions; Artificial Intelligence, Logic and Formalizing Common Sense*. Computer Science Department, Stanford University.

McCauley, L. (2007). *The Frankenstein Complex and Asimov's Three Laws*. Association of the Advancement of Artificial Intelligence.

McCauley, L. (2007). *Countering the Frankenstein Complex.* American Association for Artificial Intelligence. https://www.openculture.com/2019/03/isaac-asimov-predicts-the-future-of-civilization.html

Miller, G. (2018). *Why Altered Carbon is not about the future-nor is any other science fiction*. The Conversation.

Murphy, R. R., & Woods, D. D. (2009). *Beyond Asimov: The Three Laws of Responsible Robotics, Institute for Human and Machine Cognition*. IEEE Computer Society.

Nofz, M.P., & Vendy, P. (2002). When Computers Say It with Feeling: Communication and Synthetic Emotions in Kubrick's 2001: A Space Odyssey. *Journal of Communication Inquiry*, 25-45.

Overbye, D. (2018). 2001: A Space Odyssey' Is Still the 'Ultimate Trip'. *The New York Times*. https://www.nytimes.com/2018/05/10/science/2001-a-space-odyssey-kubrick.html

Pope, R. I. (2003). *Kubrick's Crypt, a Derrida/Deluze Monster, in 2001: A Space Odyssey*. Edinburgh University Press Journals.

Prescott, J., (2016). Why watching Westworld's robots should make us question ourselves. *The Conversation*.

Russel, S. (2015). Ethics of Artificial Intelligence. *Nature*, 415-418.

Shelley, M. (1818). *Frankenstein or The Modern Prometheus*. Harding, Mavor and Jones.

Sikdar, S. (2018). Artificial Intelligence, its impact on innovation, and the Google effect. *Clean Technologies and Environmental Policy*, *20*, No-1–No-2.

Slocombe, W. (2016). What science fiction tells us about our trouble with Artificial Intelligence, *The Conversation*.

Sloman, A. (2006). *Why Asimov's three laws of robotics are unethical.* Retrieved June 9, 2006, URL: https://www.cs.bham.ac.uk/research/projects/cogaff/misc/asimov-three-laws.html

Stahl, B. C. (2021). Organisational responses to the ethical issues of AI. *AI & Society*.

Thomsen, C. W. (1982). Robot Ethics and Robot Parody: Remarks on Isaac Asimov's *I, Robot* and Some Critical Essays and Short Stories by Stanislaw Lem. In T. P. Dunn & R. D. Erlich (Eds.), *The Mechanical God: Machines in Science Fiction* (pp. 19–26). Greenwood.

Trent, G. (2012). *The Rebirth of Mankind Homo Evolutis*. PDXdzyn.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *49*, 433–460.

Warwick, K. (2002). *I, Cyborg*. Century.

Yu, H., Shen, Z., & Miao, C. (2018). Building Ethics into Artificial Intelligence, *Computer Science, Artificial Intelligence*. Cornell University. https://arxiv.org/abs/1812.02953

KEY TERMS AND DEFINITIONS

AI Ethics: AI ethics will give an AI consciousness and intelligence to resolve conflicts, crisis, and problems which will ultimately benefit humanity involving both philosophical and technical issues. The AI should be programmed in such a way that the system can eradicate accidents and disasters as much as possible. It must deal with issues like human welfare, justice for most of the people, climate change, and defending the basic rights of human beings.

Artificial Intelligence: It is the science and engineering of constructing intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. They can very easily solve difficult and complex mathematical problems, e.g., they can play chess at a very high level or can drive a motor car on its own under human supervision.

Asimov's Three Laws: 1) A robot may not injure a human being or, through inaction, allow a human being to come to harm. 2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. 3) A robot must protect its own existence if such protection does not conflict with the First or Second Laws.

Robots: The complex machines performing a program while detecting and overcoming problems and acting towards fulfillment of the human instructions—all this makes a perfect robot. Karel Čapek, who coined the term 'robot' for machines replace human workers, related robots with slaves. In his play *Rossum's Universal Robots* (R.U.R.; 1920) he suggested that robots could replace human slave labor.

Turing's Test: In 1950 Alan Turing proposed an important condition for considering a machine to be intelligent. He emphasised that if the machines could successfully pass a test that it is a human to a knowledgeable observer then we certainly should consider it intelligent.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 29
Towards Ethical Neuromarketing 2.0 Based on Artificial Intelligence

Elodie Attié
https://orcid.org/0000-0003-3400-8927
Capgemini Engineering, France

Solène Le Bars
Altran Lab, France

Ilhem Quenel
Capgemini Engineering, France

ABSTRACT

Eighty percent of consumer behaviors and purchases rely on subconscious processes. The use of neuromarketing tools to study consumer behavior is not clear, notably regarding its practices and intentions toward consumers. This chapter aims to understand how neuromarketing can explain consumer behavior thanks to Neuromarketing 2.0 tools, how companies can manage the collected data in a responsible way and build a neuroethical charter to regulate the way companies use it. Most companies choose to not communicate about it when they use neuromarketing tools, and therefore, this chapter aims to pave the way towards solutions and recommendations and democratize its use by making Neuromarketing 2.0 more responsible and ethical.

INTRODUCTION

Between 80 and 97% of new product launches on the market fail from a lack of understanding of the market challenges and consumer behavior and needs (Mediamarketing, 2017). Neuromarketing is an area of marketing that uses neuroscience frameworks and technics, such as psychophysics or brain imaging, in order to further understand these underlying processes, and to influence consumer behaviors (Droulers

DOI: 10.4018/978-1-7998-6985-6.ch029

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

& Roullet, 2010; Morin, 2011; Nilashi et al., 2020). One of the objectives of neuromarketing is to adapt the methods and theories from neuroscience and artificial intelligence (AI) research with behavioral theories, and develop experimental designs to better understand consumer behavior and decision-making processes (Plassmann et al. 2015). Cognitive neuroscience examines how the central nervous system generates psychological processes that typically come into play when purchasing a product, such as perceptions, feelings, memory, intentions, judgment, planning, decision-making, action, behavior, etc. Therefore, neuromarketing is of prime interest for managers, in order to get an overview of the current market and its challenges. Multinational companies such as Coca Cola, McDonald's, Google, Facebook, Apple, Ebay, to name a few, have used, or use, neuromarketing tools to refine their marketing strategies, increase the frequency rate and consequent sales to ensure profits (Bayle-Tourtoulou & Badoc, 2020). However, if neuromarketing attracts managers by its opportunities and benefits, most companies that use neuromarketing tools still hide this information, stating confidential intellectual property (de Sousa, 2018). On one side, companies can increase sales and profits whereas on the other side, consumers can get better products and services, and access to better-targeted information about those products and services (Wieckowski, 2019). Neuromarketing 2.0 uses tools driven by AI, such as facial and movement coding. With neuromarketing and AI, marketing tools have become more efficient. Thus, the ethical issues raised regarding neuromarketing and AI are similar to the ones raised regarding advertising and marketing in general. The use of neuromarketing 2.0 with AI tools seems to be a wise response to address complex ethical problems, notably because they are innovative, non-invasive and they do not rely on medical data collection. Therefore, neuromarketing 2.0 tools include algorithms that are capable of collecting then interpreting data, in order to mimic cognitive decisions and predict future behaviors. For this reason, researchers and consumers may perceive neuromarketing as an immoral deviation of neuroscience, and this chapter aims at proposing an ethical way of using neuromarketing 2.0 tools. Therefore, we aim to define a neuroethical charter to protect consumers' personal data and regulate the use of neuromarketing. The main goals of the proposed chapter are the following:

1. To better understand and explain consumer behavior through conscious and automatic processes;
2. To discuss which neuromarketing 2.0 tools could be used in market research as non-intrusive and high human valued technologies;
3. To understand and change the negative perception of neuromarketing by making it more relevant, ethical and transparent to managers and consumers;
4. To build a neuroethical charter to regulate the way companies use neuromarketing through AI tools.

BACKGROUND

Marketing Research: Understanding Consumer Behavior

Researchers describe the notion of subconscious as a system or process existing in consumers' brain, and strongly determined by the environment that emit signals requiring attention, and engender emotions and memory (Guelfand, 2013). Thanks to neuromarketing, marketers can observe consumer behavior through cognitive processes such as attention, emotions and memory which all influence consumer perceptions toward the brand, and then can predict consumer behavior. Firstly, the literature distinguishes several types of attention (i.e., the ability to focus consciously or subconsciously on specific things in

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

the presence of others; Chou et al., 2010) that neuromarketing can analyze (see Figure 1). Visual attention implies an alpha activity in the occipital region. Marketers can increase consumers' attention with greater color contrasts on packaging in order to differentiate from their competitors (Burke & Leykin, 2014; Stoll et al., 2008).

Figure 1. Types of consumer attention

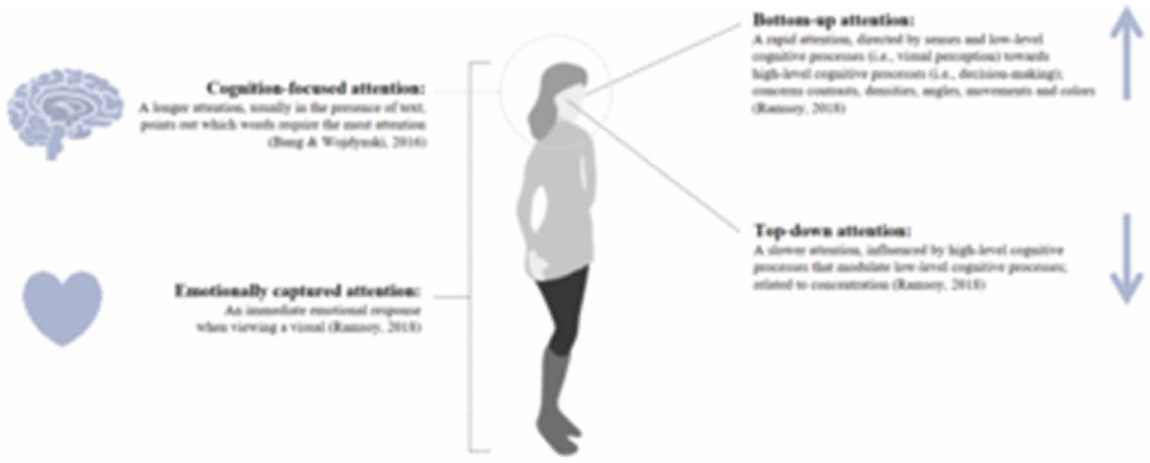


Figure 1 describes four types of attention that marketers can observe with neuromarketing 2.0 tools, notably. Eye tracking can measure the bottom-up and top-down attention. For example, Google used the technology of eye tracking to measure the influence of users' attention and emotional arousal (with pupil dilatation) while looking at new advertising formats, comparing the efficacy of overlay formats versus banners (O'Reilly, 2013). Based on this experimentation, Google could recommend to its clients that overlay formats increase attention and emotional involvement. Moreover, facial coding can analyze the emotionally captured attention. Coca Cola used facial coding to analyze quantitative ad performances (Bayle-Tourtoulou & Badoc, 2020; Dooley, 2013). To do this, Coca cola recorded consumers' facial muscles movements while they were watching their advertisings, in order to capture, analyze and understand real time emotional and cognitive states of mind induced (Dooley, 2013). Both eye tracking and facial coding can determine the cognition-focused attention.

Secondly, emotions play an essential role in consumer decision-making (Corcos & Pannequin, 2013). Generally, consumers use more emotions than rationality during a decision-making process (Murray, 2013). Emotions are characterized by an observable physical state, which can be translated into sensations (Solomon, 2017), and then into behaviors (Ciccarelli & Meyer, 2006). The type of emotion modulates specific alpha rhythms in the frontal cortex (Morris et al., 2001). Thus, a higher neuronal activity in the left frontal alpha activity implies negative emotions whereas a higher neuronal activity in the right frontal activity implies positive emotions (Vecchiato et al., 2011). Marketers usually consider two types of emotions:

1. Basic emotions: positive-negative feelings (i.e., pleasure/displeasure) associated with low to high levels of excitement (Russell et al., 1989);

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Consumption emotions: feelings linked to purchasing decisions and consumption experiences (Westbrook & Oliver, 1991).

Thirdly, memory expresses the effects of previous experiences embedded in present behaviors (Roediger, 1990). The cognitive and affective information present in advertising can create a mental representation that enhance long-term memory (Plassmann et al., 2015). This process of retention of the information and events takes places in the left frontal region. There is a diminution of the alpha activity which reflects the encoding phase in long-term memory (Langleben et al., 2009; Rossiter et al., 2001).

Marketing Evolution: From Marketing Research to Neuromarketing 2.0

Companies can use marketing research tools to adjust a marketing mix strategy, target their consumers, reformulate their offer and sales strategies, and generate new product concepts to increase business profits, among other marketing challenges (Guelfand, 2013). Traditional marketing research attempts to catch consumers' intentions and behaviors, by directly asking consumers how they perceive a brand, product or service. Consumers translate their unconscious thoughts into language, sometimes allowing brands to grasp subconscious intentions (Guelfand, 2013). However, traditional market research remains subjective and entails many biases due to confounding factors such as social desirability or environmental influence (Legardinier, 2013). Neuromarketing should reduce the uncertainty to understand and explain consumer behavior (Harrell, 2019). In order to limit these biases, neuromarketing employs neuroscience tools to get physiological or alternative behavioral measures and capture consumers' subconscious response to marketing strategies (Oullier, 2012) (see Figure 2).

Figure 2 shows that neuromarketing techniques can evaluate:

1. Brain electrical and chemical modifications, thanks to traditional medical tools like fMRI (functional magnetic resonance imaging), EEG (electroencephalogram), or MEG (magnetoencephalography);
2. Physiological activities, with other traditional medical tools like ECG (electrocardiograms), GSR (galvanic skin responses), or facial electromyograms, and AI tools like eye tracking and facial coding;
3. Behaviors and attitudes, with the implicit response test, response times, and indoor GPS.

For example, McDonald's used fMRI, with a neuromarketing company based in London called Neuro Sense, to measure the brain activation of the reward circuit (i.e., the neural system responsible for motivation, desire and positive reinforcement) with participants who were watching their ads and measure their effectiveness (Bayle-Tourtoulou & Badoc, 2020).

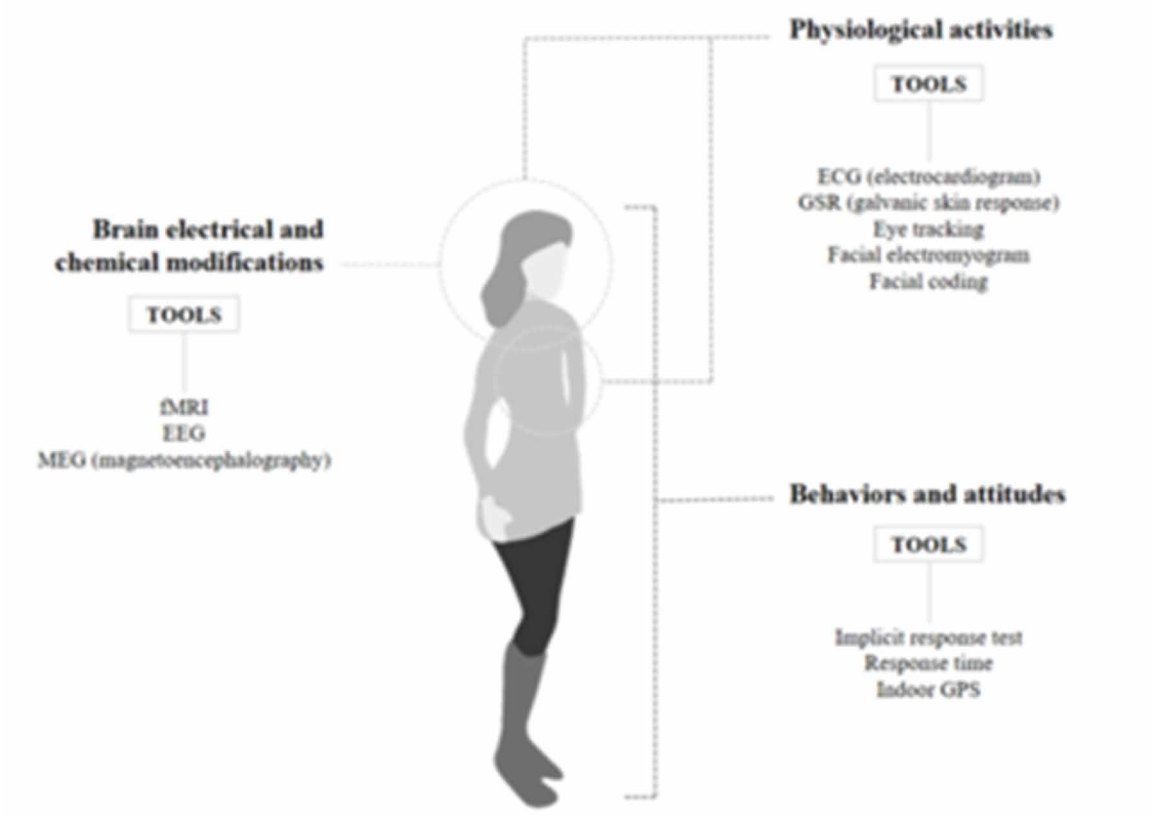
SETTING THE STAGE

Understanding how the Brain Influences Consumer Behaviors Through Neuromarketing Lens

Each region of the brain that activates implies specific functions and in turn specific behaviors. The nucleus accumbens is the region of positive emotions that activates when consumers perceive advan-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

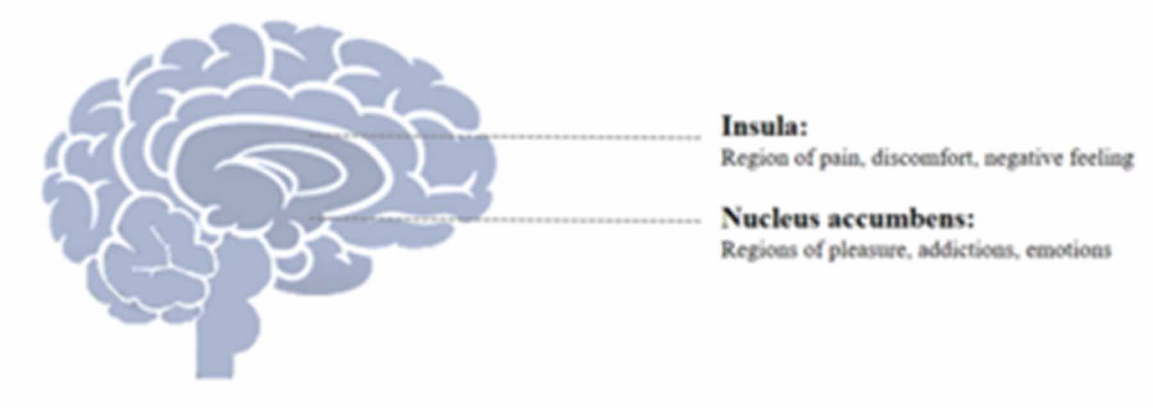
Figure 2. Consumer indicators and neuromarketing tools



tages and gains –explaining impulsive purchases for example (Corcos & Pannequin, 2013; Knutson et al., 2007). However, when consumers perceive expensive prices, the insula activates, then frontal areas activate, in order to use rationality and judgment –the purchase decision is slower and more cautious (Corcos & Pannequin, 2013; Knutson et al., 2007). The medial prefrontal cortex evaluates the trade-off between the advantages of the product or service –perceived through the nucleus accumbens– and the consequences of the choice –perceived through the insula (Knutson et al., 2007). Thus, companies can measure unexpected feelings toward a product or service, and predict behaviors according to the activation of different brain areas (Damasio, 1994; Bechara et al., 1997; Knutson et al., 2007). For example, watching an attractive advertising or product induces the activation of a more specific brain area involved in the reward circuitry: the nucleus accumbens (Knutson et al., 2007) –linked to the anticipation of pleasure (Carter et al., 2009). In contrast, the perception of a high price can curb the product purchase and this phenomenon results from another brain area, the insula –linked to the anticipation of pain (Knutson et al., 2007). Figure 3 presents the differentiation between the insula and nucleus accumbens regions of the brain.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Regions of the brain activated with positive and negative emotions



How Neuromarketing can Influence Consumer Behaviors and Intentions to Purchase

Companies usually use fMRI to study the influence of packaging, ads, logos, or slogans on the brain (Bayle-Tourtoulou & Badoc, 2020). Neuromarketing thus enables to redefine packaging and product conception and advertising, and thus, promote specific behaviors (Harrell, 2019). Neuromarketing also embraces computational models of decision-making, big data and AI algorithms to better predict consumers' behaviors, becoming a neuromarketing 2.0 (Hegazy, 2019; Stoicescu, 2016). Indeed, big data and AI tools allow companies to survey and select relevant information from their consumer target, apply analytical tools, build solutions from market predictions, and find relationships between factors that may influence consumer behavior (Phillips-Wren, 2012). Neuromarketing 2.0 can lead companies through AI techniques in order to study consumers' subconscious thoughts, which represent inaccessible contents and knowledge about decision-making and actions (Guelfand, 2013). Consumer behavior evolves with the activity of neuromodulators –brain hormones such as testosterone, cortisol and oxytocin– and neurotransmitters –chemical messengers. For example, dosing artificial testosterone to men increases their preference for higher social status and luxury brands, compared to brands of similar perceived quality but synonyms of lower perceived social status (Nave et al., 2018). Thus, neuromarketing tools allow managers to understand how to activate the right lever in consumers' brains to create an action toward the brand, products and services (de Sousa, 2018). Researchers scanned consumers' brains while they tasted three glasses of wines with three different prices: even though the glasses contained the same wine, each glass of wine activated a different area in consumers' brains, highlighting a preference for the most expensive wine (Schmidt et al., 2017). Moreover, subliminal advertising (i.e., presenting a particular stimulus below the threshold of conscious perception) constitutes a representative example of how neuromarketing can use cognitive neuroscience methods (Smarandescu & Shimp, 2015; Sofi et al., 2018). Brands can display a subliminal message (i.e., word, image, emoji, logo or jingle) to influence intentions and behaviors toward the brand (Hsu et al., 2020; Smarandescu & Shimp, 2015; Sofi et al., 2018). This technique influences consumers' perceptions and cognition, at an automatic level, and interferes with consumers' endogenous goals and desires regarding a product or behavior. However, there is no evidence that subliminal advertising is effective in real market conditions.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

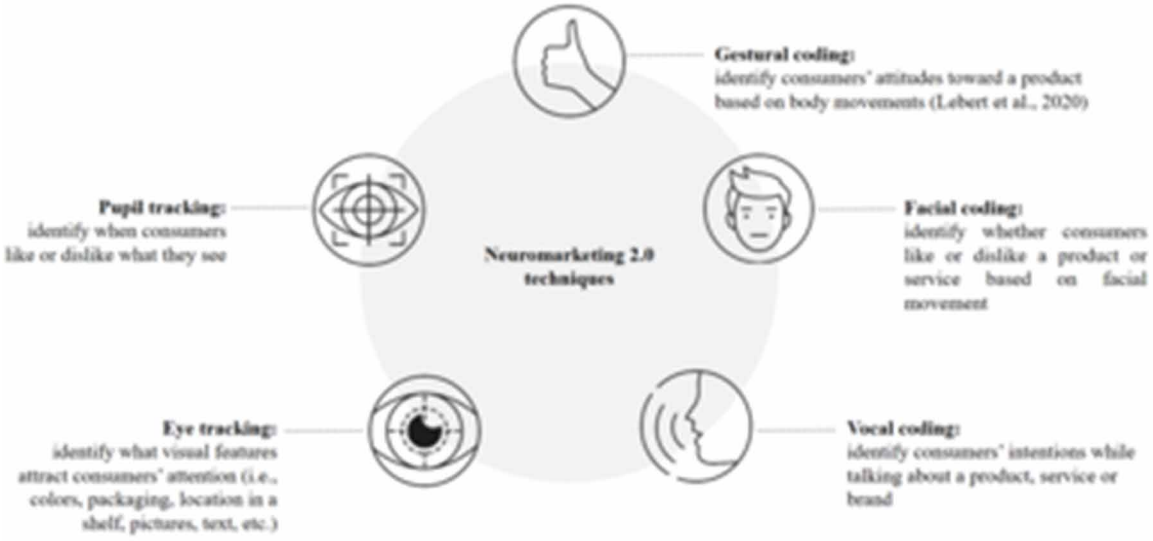
CASE DESCRIPTION

Neuromarketing 2.0: Technology Components and Concerns

Technology components

According to the literature, neuromarketing presents some advantages. The main drivers for neuromarketing 2.0 adoption are economic, such as cost and time reduction, an improved performance, greater customer satisfaction, and more accurate forecasting and decision-making (Cubric, 2020). AI can now exceed 87% of sensitivity, 91% of specificity and 96% of predictions, thus offering security, efficiency and equity (Abràmoff et al., 2020). If neuromarketing techniques are well used, the data can bring out a huge competitive advantage to companies, defining predictive models based on feeling detection for example (de Bruyn et al., 2020). Specifically, Figure 4 presents the benefits and risks of neuromarketing 2.0 tools that seem to be the most relevant to this research, such as gestural coding, facial coding, vocal coding, eye tracking and pupil tracking.

Figure 4. Neuromarketing 2.0 techniques



All these neuromarketing 2.0 techniques (Figure 4) enable marketers to predict behaviors and intentions to purchase. AI can identify key points from facial images based on facial activities and muscles, and learn the visual patterns of facial movements according to emotional labels (Calvo & D'Mello, 2010). Moreover, there are important relationships between voice parameters and emotions or behaviors (Banse & Scherer, 1996; Johnstone & Scherer, 2000). Regarding the eye tracking method, there are correlations between the time spent looking at the advertising and preferences (Maughan et al., 2007; Orquin & Loose, 2013). Finally, the pupil size increases during both emotionally negative and positive stimuli when compared to neutral stimuli (Partala & Surakka, 2003).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Technology Concerns

The main technology issues with neuromarketing 2.0 concern its technical (i.e., data training and data management) and social features (i.e., dependence on technology, technology trust) (Cubric, 2020). Firstly, there is a significant cost of implementation and maintenance, and if necessary, of support infrastructures: the creation of algorithms able to capture, explain and predict behaviors requires significant costs linked to the specific machines implied and people to hire. Maintenance also involves additional costs. Secondly, neuromarketing 2.0 has a limited applicability to specific issues: the relationships highlighted are often correlational rather than causal (Ma & Sun, 2020). This lack of causal capacity makes predictions unstable when change occurs, limiting their use in marketing (Ma & Sun, 2020). For example, one of the limits of eye tracking is the lack of knowledge to understand why consumers look at a particular spot on an advertising (Lever et al., 2019). Thirdly, technology itself does not have emotions and moral values. Neuromarketing 2.0 executes its program and cannot judge what is right or wrong like humans do. Yet, human bias also imply technology bias (Cubric, 2020). Indeed, people can be biased and judgmental, leading to biased and judgmental technologies that try to explain consumer behavior (de Bruyn et al., 2020). Therefore, cybersecurity is an important asset of AI to ensure a safe and positive evolution of technology in society (Lin et al., 2011). Finally, despite some obvious benefits of using neuromarketing 2.0 to develop efficient marketing strategies, this technique also raises some important ethical questions. First, neuroimaging technics can be invasive (Oullier, 2012) –except for GSR, eye tracking, or response time measurement– and even though some neuromarketing technics are non-invasive, their safety and reliability still represent ethical issues (Claeys & Vialatte, 2012). In France, only scientific and medical research can use fMRI and other hospital scanners. However, other countries like United Kingdom, Belgium or the United States are more open about the development and use of neuromarketing.

Neuromarketing 2.0: Consumer Privacy Concerns

Neuromarketing 2.0 presents security and ethical issues. The way the technology tracks and collects personal data can be intrusive, arousing privacy concerns (Awad & Krishnan, 2006; Hong & Thong, 2013; Phelps et al., 2001). Privacy concerns is the extent to which consumers are concerned about the flow of their information through technologies (Phelps et al., 2000). More specifically, consumers create beliefs and trust toward the use of the technology and privacy management (Chaudhuri & Holbrook, 2001; Phelps et al., 2000; Shin, 2010). For consumers, their purchases can be known from companies but their thoughts and feelings should remain private to them (Wieckowski, 2019). Consumers might perceive companies that use neuromarketing tools as untrustworthy since they can invade their private thoughts and feelings, sometimes leading to consumer backlash against the brand (Wieckowski, 2019). Therefore, regardless of the neuromarketing and market research methodology used (i.e., neuroscience tools, data mining based on activity and purchase history, consumer surveys, etc.), consumers may perceive reticence due to privacy concerns and a violation of their right to privacy (Wieckowski, 2019). However, the benefits of neuromarketing 2.0 (i.e., personalization of products and services) can compensate the loss of privacy (Xu et al., 2011). Moreover, there are few ethical reviews of neuromarketing 2.0, mainly due to the novelty and complexity of this technique (Abràmoff et al., 2020). Thus, companies have to collect data according to each country's regulatory framework, in a way that is transparent regarding the purpose and the extent of use of the data (Abràmoff et al., 2020). In Europe, since May 2018, a recent law called GDPR (General Data Protection Regulation) regulates the use of data and ensures the protection

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of data, with a unified set of rules for all organizations that process personal data coming from Europe. The GDPR improves data management by giving more visibility and control to European citizens over the management of their data (i.e., what data, what purpose, how long). If European organizations, or worldwide organizations collecting European data, do not follow this regulation, they risk sanctions, such as 4% of their annual worldwide turnover or 20 million euros (the greater of the two), and additional damages to the parties concerned (see cnil.fr website for updates).

Neuromarketing 2.0: Management and Organizational Concerns

When neuromarketing 2.0 tools are well designed and used, it can deepen the understanding of consumers and business, point out coming issues, and suggest automated decision support (Ma & Sun, 2020). Therefore, the cooperation between humans and technology predicts important changes in organizations, with an evolution of organizational behaviors, structures, workflows, ethics and work environments (Demir et al., 2019). It becomes relevant to focus on the management, organizational and social aspects of neuromarketing 2.0, such as the factors of acceptance (Cubric, 2020), the need for education and training, and psychological, social and ethical issues, among other issues (Demir et al., 2019). According to Lin et al. (2011), the main management challenges and recommendations imply:

1. **Performance:** Companies must be able to create a neuromarketing 2.0 that can do minimal distinctions in gestures or speech and take into account relevant and salient safety information;
2. **Ethical regulation:** Companies should define a code of ethics to regulate neuromarketing 2.0. and be clear about their rights and responsibilities, as well as consumers';
3. **A budget:** Neuromarketing 2.0 implies an economic impact of AI tools and robotics, and companies should be able to estimate its expected costs and benefits on the long-term.

Even though the development of technologies can threatens the survival of certain jobs, people remain irreplaceable in the assessment of subjective and qualitative issues (i.e., intuitive capacities, norms, intangible interests and other complex social and contextual factors like decision-making) (Demir et al., 2019; Jarrahi, 2018). Thus, neuromarketing 2.0 can extend human capabilities rather than replace human decision-making (Jarrahi, 2018).

CURRENT CHALLENGES

Data Management: How to Manage Meaningful Information

Neuromarketing 2.0 imply a collection and management of all kinds of data (i.e., pupil dilatation, facial expressions, different position in a store, etc.). There should be a strategic and structured vision of the collected data. Data management can increase the productivity and overall performance of the company thanks to the quality of data, compliance with requirements agreed by the user, efficiency gains, access and security of the data (Soler et al., 2016). Data management implies standards and tools, ensuring the accessibility, reliability and security of data. Therefore, data management requires technical expertise regarding the data programming and quality, and data transmission, extraction and use –implying time and money to train employees to learn how to manage the data (Olajide, 2019). The main goal of data

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

management is to transform data into valuable information for business and strategies. Data management teams intervene at the following levels:

1. **Data quality:** Data quality is the basis of data management by ensuring accuracy, consistency, and traceability. Non-quality data represent a significant risk and additional cost for the company since incorrect or not updated data can distort customer knowledge and decrease the reactivity of decisions. Thus, companies should have clear objectives, good material and know how to analyze the data. For example, studying the attention process of consumers requires a good eye tracking material and people able to analyze this type of data. Moreover, companies should constantly collect and update the data in order to keep up with the evolution of data innovations. With data quality, neuromarketing can, and should, guide marketers and managers in decision-making and action, through anticipation, predictions and smart recommendations.
2. **Data security:** According to the collected data, security can become a very important issue. Indeed, neuromarketing 2.0 can sometimes track very personal data (i.e., images or videos of faces and body movements) that consumers do not want to spread. Therefore, it is important for them to know how the collected data are stored and used.
3. **Data repository:** Data management implies the need of a master data management. Master data represents the strategic information that marketers need to collect to understand consumer perceptions. It could be information about customers, products, resources, market challenges, etc.
4. **Compliance with legal obligations:** Legal obligations mainly come from the GDPR to encourages an ethical use and traceability of the data within companies, through five principles:
    - *Accountability*: the obligation to implement internal mechanisms and procedures to demonstrate the compliance with data protection rules;
    - *Data protection by design*: the implementation of technical and organizational measures, at the earliest stages of the design of the processing operations, to protect privacy by following data protection principles –it usually implies the use of anonymization (i.e., replacing personal identifiable material with artificial identifiers) and the use of encryption (i.e., encoding messages so that only those authorized can access the data);
    - *Data protection by default*: the process of ensuring the highest privacy protection –it implies settings like a limited accessibility, a short period storage, and only the data necessary for companies' use is processed;
    - *Data protection officer*: the data protection officer is responsible for the protection and compliance of the data –leading to the development of new jobs in the data sector;
    - *Data impact assessment*: a detailed description of the processing (i.e., technical and operational aspects), a legal assessment of the fundamental principles and rights (i.e., purpose, what data, retention periods, information and rights of individuals), and a technical study of data security risks (i.e., confidentiality, integrity and availability).
5. **Infrastructures:** The main challenge of companies is to optimize neuromarketing infrastructures. Data management structures must meet both the accessibility needs of neuromarketing and business, as well as the regulations and security constraints.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Neuroethics: How to use Neuromarketing 2.0

Consumers may perceive some doubts and fears about neuromarketing intentions and techniques (Attié et al., 2021). Consumers perceive neuromarketing as a way to influence them, through persuasion and manipulation, with few concerns about the benefits for consumers and data protection (Flores et al., 2014). Yet, some ethical principles in neuromarketing must be followed, such as asking for participants' consent regarding the use of the collected data –which do not differentiate much from traditional marketing (Wieckowski, 2019). Neuromarketing 2.0 leads to technologies that have not reached yet all the capabilities of the human brain when it comes to detect emotions, be empathetic, have an intuition; its values and morality depends on how people decide to build it (Baillargeon, 2018). Thus, one challenge is to understand the impact of AI on neuromarketing, in order to highlight the best practices and benefits for researchers, managers and consumers. A way to face ethical issues due to neuromarketing could be to create and adopt an ethical and responsible way of treating the data (Sultana & Nangunoori, 2008). For this reason, this chapter aims to build a neuroethic charter to respond to deontology aspects raised by the use of AI tools in neuromarketing 2.0. Neuroethics focus on both theoretical and practical aspects of implementing AI into neuromarketing challenges since it examines how the knowledge of the brain can explain consumers' identity, consciousness and intentions (Evers, 2014). The theoretical foundations of neuroethics should enable companies to address ethical challenges linked to neuromarketing (Evers, 2014).

SOLUTIONS AND RECOMMENDATIONS

Ethics and Transparency: Improving the Acceptance of Neuromarketing 2.0

Neuromarketing enables to observe consumers' conscious thinking and identify automatic reactions. By combining both neuromarketing and traditional marketing research methods, companies show an interest to understand and invest in consumers to respond to their needs (Sultana & Nangunoori, 2008). However, employees and consumers may perceive neuromarketing as immoral due to this perception of finding ways to manipulate consumers (Charmettant, 2018). Managers could thus change the way consumers perceive neuromarketing and transform its dark sides into a clarified explanation of this marketing tool, through communication, transparency, and clear guidelines and infrastructures –both internally and externally (Attié et al., 2021; Courbet & Benoit, 2013). Thus, communication should demonstrate the benefits of neuromarketing techniques, as some research has shown its futility (Sultana & Nangunoori, 2008). Indeed, explaining neuromarketing processes and strategies, and how they influence consumers, could mitigate the unconscious effects and doubts (Courbet & Benoit, 2013). Moreover, consumers expect higher prices if companies use neuromarketing tools since it should increase companies' budgets in marketing (Attié et al., 2021). Yet, higher prices are not a main brake (Burke & Leykin, 2014) if it is justified by higher quality, positive innovative brand image, positive emotions (Berger & Milkman, 2012), and a new experience to enjoy that neuromarketing could create (Attié et al., 2021). In addition, implementing neuromarketing 2.0 means that employees should accept this technology as well, developing a gain in skills and a synergy resulting from the appropriation of these new tools (Quéant, 2019). Thus, companies should slowly prepare and integrate this change into their organization and strategy (Quéant, 2019). Finally, it is important to find ethical ways to use neuromarketing tools. Researchers and marketers can understand how to push people to perform more ethical, responsible and positive behaviors thanks

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

to the use of nudges for example (Singler, 2015). The main goal of nudges is to offer a better option that maximizes health, well-being and wealth (Lin et al., 2017). However, nudges should preserve consumers' rights to choose the option they prefer (i.e., increase physical activity, decrease alcohol or smoking habits, etc.) without forcing them (Lin et al., 2017). For example, research demonstrated that smokers exposed to the smell of cigarettes mixed with rotten eggs during the phase 2 of their sleep –when the body prepares for deep sleep– would then reduce their smoking habits for several days, but smokers must be willing to stop cigarettes first (Hurst, 2015). Neuromarketing can thus predict intentions and find ways to transform these intentions into action.

Ethics and Responsibility: A Neuroethical Charter for Neuromarketing 2.0

Neuromarketing 2.0 implies collecting then analyzing and using sometimes sensitive and personal data. Marketers should thus implement and promote ethical principles linked to their use of neuromarketing. Thus, it is necessary to develop reliable and trustable neuromarketing 2.0 tools regarding the following aspects: ethical considerations, ethical practices, responsible data management, and consumer rights, according to the field of intervention.

Ethical Consideration

Companies should predict the dangers related to neuromarketing studies. For example, managers should implement safeguards policies and consider the establishment of mechanisms that promote integrity, autonomy, safety, privacy, non-discrimination and the dignity of people in the short and long term. In addition, there should be a priority given to safety assessment in the development and use of neuromarketing tools: companies should collaborate with researchers, health professionals, engineers, employees, consumers, as well as private actors and public authorities in order to prioritize research rather than business matters. This communication should address the ethical, legal and social issues, and opportunities related to advances in brain research and neuromarketing 2.0.

Ethical Practices

Companies should take into account consumers' concerns about neuromarketing 2.0 intentions and tools, from the planning and design step of the technology development to the implementation of the strategy on the market. There should be an establishment of standards and best practices of the technical ethical aspects of innovation in neuromarketing tools. Consumers should have access to sufficient and clear information about the safety, quality and efficacy of new neuromarketing 2.0 tools and procedures. In addition, companies should establish business practices that encourage accountability (i.e., social responsibility), transparency (i.e., expose the interest at stake and promote clear communications about neuromarketing tools and objectives), integrity (i.e., foster debates between the different actors involved), reliability (i.e., standards and practices that increase trust and confidence), responsiveness, and security. Therefore, there should be a willingness to encourage the further development of ethical guidelines and best practices based on the rigor and replicability of neuromarketing 2.0 tools. Finally, companies could name advisory boards specialized on ethical matters to verify compliance with ethical and legal principles, and its application.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Responsible Data Management

It is necessary to provide people and research stakeholders with clear information and education on the collection, storage, analysis and potential use of personal brain data collected. Companies should be transparent and communicate about their strategy in order to let both employees and consumers know the concept of neuromarketing 2.0, its purposes and consequences. Moreover, companies doing neuromarketing should respect existing laws regarding data use (i.e., GDPR principles in Europe) and anonymize the collected data to protect consumers' information, among the other recommendations.

Consumer Rights

Companies should avoid potential harmful effects and give due consideration to human rights and societal values, and more specifically, emphasize respect for consumer privacy, cognitive freedom and individual autonomy. Consumers should always be able to give an informed, full and free consent about their data use, in order to protect their autonomy and privacy. This consent should include an easy access to the modification and deletion of their personal data, as well as clear policies about legal proceedings, confidentiality, and privacy security. Therefore, companies need to ensure the traceability of the collected data through neuromarketing 2.0 tools. More specifically, managers should encourage efforts to raise awareness of potential existing unethical neuromarketing intentions and techniques and, if necessary, adapt to existing legislation and regulations.

Fields of Intervention

If neuromarketing can bring positive outcomes (i.e., well-being, health benefits), regulations should be less strict than if the main purpose is only to increase business profits. In some countries, there are some laws that regulate non-interventional research (i.e., human observation, data collection, and collection by extra-corporeal non-invasive sensors, audio, video and image recordings, collection of electrophysiological data, anthropometric measurements without invasive intervention, interviews, observations, tests and questionnaires).

CONCLUSION

Marketers need to gain consumer acceptance regarding any marketing techniques they consider using (du Pre Gauntt, 2008; Merisavo et al., 2007; Ranchhod, 2007). In 2015, most people talking to an AI bot named Eugene Goostman believed they were talking to a real person. Therefore, smart robots can offer great opportunities in sales services by building relationships with customers, while analyzing and predicting their behavior. This chapter aims at highlighting and discussing reliable and trustable neuromarketing 2.0 tools to detect and translate consumer attention, emotions, intentions and behaviors. Thanks to neuromarketing, companies should get access to relevant tools to create better ads for the greater good of everyone, by encouraging positive behaviors or even citizenship through nudges (Falk et al., 2015; Wieckowski, 2019). Another contribution was to define a neuroethical charter to favor transparency about the technologies used and companies' intentions, and to develop ethical and trustworthy data management guidelines. This chapter aims at encouraging research on the ethical, legal and societal

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

dimensions of neuromarketing 2.0. If used right, neuromarketing can become a catalyst for positive change in the coming years. Thus, this chapter should give an understanding of how neuromarketing 2.0 works with AI tools, how to use these tools, in order to improve marketing strategies and relationships between companies and consumers. Finally, almost all companies protect their intellectual property when they use neuromarketing and do not communicate much about it, but this issue could be solved by rewarding the firms that decide to be transparent and willing to be peer reviewed for example, leading the way toward more ethical business and practices (Wieckowski, 2019).

ACKNOWLEDGMENT

This research was supported by Capgemini Engineering.

REFERENCES

Abràmoff, M. D., Tobey, D., & Char, D. S. (2020). Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. *American Journal of Ophthalmology*, *214*, 134–142. doi:10.1016/j.ajo.2020.02.022 PMID:32171769

Attié, E., Gourdou, C., & Guédon, E. (2021). *Understanding consumer acceptance of neuromarketing* Paper presented at *International Marketing Trends Conference*, Venice.

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *Management Information Systems Quarterly*, *30*(1), 13–28. doi:10.2307/25148715

Baillargeon, N. (2018). L'éthique des voitures autonomes [The ethics of autonomous cars]. *Québec Science Magazine*. Accessed from: https://www.quebecscience.qc.ca/normand-baillargeon/ethique-des-voitures-autonomes/

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. doi:10.1037/0022-3514.70.3.614 PMID:8851745

Bayle-Tourtoulou, A.-S., & Badoc, M. (2020). *The neuro-consumer: Adapting communication and marketing strategies for the consumer*. Routledge. doi:10.4324/9781003019978

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293–1295. doi:10.1126cience.275.5304.1293 PMID:9036851

Berger, J., & Milkman, K. (2012). What makes online content viral? *JMR, Journal of Marketing Research*, *49*(2), 192–205. doi:10.1509/jmr.10.0353

Burke, R., & Leykin, A. (2014). *Identifying the Drivers of Shopper Attention*. Engagement, and Purchase., doi:10.1108/S1548-643520140000011006

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. doi:10.1109/T-AFFC.2010.1

Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Frontiers in Behavioral Neuroscience*, *3*, 21. doi:10.3389/neuro.08.021.2009 PMID:19753142

Charmetant, E. (2018). La neuroéthique: Histoire, actualité et prospective [Neuroethics: history, present and future]. *Transversalites*, *146*(3), 45–58. doi:10.3917/trans.146.0045

Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing*, *65*(2), 81–93. doi:10.1509/jmkg.65.2.81.18255

Chou, K. L., Amick, M. M., Brandt, J., Camicioli, R., Frei, K., Gitelman, D., Goldman, J., Growdon, J., Hurtig, H. I., Levin, B., Litvan, I., Marsh, L., Simuni, T., Tröster, A. I., & Uc, E. Y. (2010). A Recommended Scale for Cognitive Screening in Clinical Trials of Parkinson's Disease. *Movement Disorders*, *25*(15), 2501–2507. doi:10.1002/mds.23362 PMID:20878991

Ciccarelli, S. K., & Meyer, G. E. (2006). *Psychology*. Prentice Hall.

Claeys, A., & Vialatte, J.-S. (2012). Office parlementaire d'évaluation des choix scientifiques et technologiques. L'impact et les enjeux des nouvelles technologies d'exploration et de thérapie du cerveau [Parliamentary Office for the Evaluation of Scientific and Technological Choices. The impact and challenges of new technologies for brain exploration and therapy]. *Rapport par Claeys et Vialatte, députés. Paris*. *Assemblée Nationales et Sénat*, *2012*, 25.

Corcos, A., & Pannequin, F. (2013). Neuroeconomics, decision-making and rationality. *Économie et institutions, 16*. . doi:10.4000/ei.74

Courbet, D., & Benoit, D. (2013). Neurosciences au service de la communication commerciale: manipulation et éthique - Une critique du neuromarketing [Neuroscience at the service of commercial communication: manipulation and ethics - A Critique of Neuromarketing]. *Études de communication - Langages, information, médiations*. *Université de Lille*, *40*, 28–42.

Cubric, M. (2020). Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study. *Technology in Society*, *62*, 101257. Advance online publication. doi:10.1016/j.techsoc.2020.101257

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. G.P. Putmans & Sons.

De Bruyn, A., Viswanathan, V., Beh, Y. S., Brock, J. K. U., & von Wangenheim, F. (2020). Artificial Intelligence and Marketing: Pitfalls and Opportunities. *Journal of Interactive Marketing*, *51*, 91–105. doi:10.1016/j.intmar.2020.04.007

De Sousa, J. (2018). *The Challenge of Neuromarketing*. IGI Global. doi:10.4018/978-1-5225-4834-8.ch002

Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and human-robot co-working. *Procedia Computer Science*, *158*, 688–695. doi:10.1016/j.procs.2019.09.104

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dooley, R. (2013). *Neuromarketing: For Coke*. It's the Real Thing.

Droulers, O., & Roullet, B. (2010). *Neuromarketing. Le marketing revisité par les neurosciences du consommateur* [Neuromarketing. Marketing revisited by the neurosciences of the consumer?]. Dunod.

Du Pre Gauntt, J. (2008). Mobile advertising: After the growing pains. New York, NY: eMarketer.

Evers, K. (2014). *Quand la neuroscience se réveille* [When neuroscience wakes up]. Upsala Universiteit. Accessed from: https://www.crb.uu.se/digitalAssets/445/c_445294-l_1-k_k-evers-neuroethique.pdf

Falk, E. B., Berkman, E. T., Whalen, D., & Lieberman, M. D. (2015). Neural Activity During Health Messaging Predicts Reductions in Smoking Above and Beyond Self-Report. *Health Psychology*. PMID:21261410

Flores, J., Baruca, A., & Saldivar, R. (2014). Is neuromarketing ethical? Consumers say yes. Consumers say no. Journal of Legal. *Ethical and Regulatory Issues*, *17*(2), 77–91.

Guelfand, G. (2013). *Les études qualitatives: Fondamentaux, méthodes, analyse, techniques* [Qualitative studies: Fundamentals, methods, analysis, techniques]. EMS Editions. doi:10.3917/ems.guelf.2013.01

Harrell, E. (2019). Neuromarketing: What you need to know. *Harvard Business Review*. Accessed from: https://hbr.org/2019/01/neuromarketing-what-you-need-to-know

Hegazy, I. M. (2019). The effect of political neuromarketing 2.0 on election outcomes. *Review of Economics and Political Science*.

Hong, W., & Thong, J. (2013). Internet privacy concerns: An integrated conceptualisation and four empirical studies. *Management Information Systems Quarterly*, *37*(1), 275–298. doi:10.25300/MISQ/2013/37.1.12

Hsu, L., & Chen, Y.-J. (2020). Neuromarketing, subliminal advertising, and hotel selection: An EEG study. *Australasian Marketing Journal*, *28*(4), 200–208. Advance online publication. doi:10.1016/j.ausmj.2020.04.009

Hurst, J. (2015). Assessing kindship through scent – the mouse model. *Chemical Senses*, *40*(3), 211–297. doi:10.1093/chemse/bju073

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision-making. *Business Horizons*, *61*(4), 577–586. doi:10.1016/j.bushor.2018.03.007

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In *Handbook of Emotions* (2nd ed., pp. 220–235). Guilford Press.

Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, *53*(1), 147–156. doi:10.1016/j.neuron.2006.11.010 PMID:17196537

Langleben, D. D., Loughead, J. W., Ruparel, K., Hakun, J. G., Bush-Winokur, S., Holloway, M. B., Strasser, A. A., Cappella, J. N., & Lerman, C. (2009). Reduced prefrontal and temporal processing and recall of high "sensation value" ads. *NeuroImage*, *46*(1), 219–225. doi:10.1016/j.neuroimage.2008.12.062 PMID:19457412

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Legardinier, A. (2013). *Comment limiter les biais liés au choix des échelles de mesure dans les études marketing?* [How to limit the biases linked to the choice of measurement scales in marketing studies?]. Gestion et management, Dumas.

Lever, M. W., Shen, Y., & Joppe, M. (2019). Reading travel guidebooks: Readership typologies using eye-tracking technology. *Journal of Destination Marketing & Management*, *14*, 1–13. doi:10.1016/j.jdmm.2019.100368

Lin, P., Abney, K., & Bekey, G. (2011). Robot Ethics: Mapping the Issues for a Mechanized World. *Artificial Intelligence*, *175*(5-6), 942–949. doi:10.1016/j.artint.2010.11.026

Lin, Y., Osman, M., & Ashcroft, R. (2017). Nudge: Concept, effectiveness, and ethics. *Basic and Applied Social Psychology*, *39*(6), 293–306. doi:10.1080/01973533.2017.1356304

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing – Connecting computing power to human insights. *International Journal of Research in Marketing*, *37*(3), 481–504. doi:10.1016/j.ijresmar.2020.04.005

Maughan, L., Gutnikov, S., & Stevens, R. (2007). Like more, look more. Look more, like more: The evidence from eye tracking. *Journal of Brand Management*, *14*(4), 335–342. doi:10.1057/palgrave.bm.2550074

Mediamarketing. (2017). *Le chemin passe par le neuromarketing* [The path goes through neuromarketing]. Accessed from: https://www.mediamarketing.ma/article/GHBGBGH/le_chemin_passe_par_le_neuromarketing.html

Merisavo, M., Kajalo, S., Karjaluoto, H., Virtanen, V., Salmenkivi, S., Raulas, M., & Leppäniemi, M. (2007). An empirical study of the drivers of consumer acceptance of mobile advertising. *Journal of Interactive Advertising*, *7*(2), 1–19. doi:10.1080/15252019.2007.10722130

Morris, P. L., Hopwood, M., Whelan, G., Gardiner, J., & Drummond, E. (2001). Naltrexone for alcohol dependence: A randomized controlled trial. *Addiction (Abingdon, England)*, *11*(15), 65–73. doi:10.1046/j.1360-0443.2001.961115654.x PMID:11784454

Murray, P. N. (2013). *Inside the Consumer Mind How Emotions Influence What We Buy The emotional core of consumer decision making*. Retrieved from https://www.psychologytoday.com/blog/inside-the-consumermind/201302/how-emotions-influence-what-we-buy

Nave, G., Nadler, A., Dubois, D., Zava, D., Camerer, C., & Plassmann, H. (2018). Single-dose testosterone administration increases men's preference for status goods. *Nature Communications*, *9*(1), 24–33. doi:10.103841467-018-04923-0 PMID:29970895

Nilashi, M., Samad, S., Ahmadi, N., Ahani, A., Abumalloh, R. A., Asadi, S., & Yadegaridehkordi, E. (2020). Neuromarketing: A review of research and implications for marketing. *Journal of Soft Computer Decision Support System*, *7*(2), 23–31.

O'Reilly, L. (2013). *Google patents 'pay-per-gaze' eye-tracking ad technology*. Accessed from: https://www.marketingweek.com/google-patents-pay-per-gaze-eye-tracking-ad-technology/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Olajide, V. (2019). *Introduction to Data Management*. Accessed from: https://www.researchgate.net/publication/335062781

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206. doi:10.1016/j.actpsy.2013.06.003 PMID:23845447

Oullier, O. (2012). Le cerveau et la loi: analyse de l'émergence du neurodroit [The Brain and the Law: Analysis of the Emergence of Neurolaw]. Centre d'Analyse Stratégique.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1-2), 185–198. doi:10.1016/S1071-5819(03)00017-X

Phelps, J. E., D'Souza, G., & Nowak, G. J. (2001). Antecedents and consequences of consumer privacy concerns: An empirical investigation. *Journal of Interactive Marketing*, *15*(4), 2–17. doi:10.1002/dir.1019

Phelps, J. E., Nowak, G. J., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing*, *19*(1), 27–41. doi:10.1509/jppm.19.1.27.16941

Phillips-Wren, G. (2012). AI tools in decision making support systems: A review. *International Journal of Artificial Intelligence Tools*, *21*(2), 1240005. doi:10.1142/S0218213012400052

Plassmann, H., Ramsoy, T. Z., & Milosavljevic, M. (2015). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology*, *22*(1), 18–36. doi:10.1016/j.jcps.2011.11.010

Quéant, P. (2019). IA et puissance aérospatiale, implications pour l'Armée de l'air [AI and Aerospace Power, Implications for the Air Force]. *Revue Défense Nationale*, *820*(5), 117–122. doi:10.3917/rdna.820.0117

Ranchhod, A. (2007). Developing mobile marketing strategies. *International Journal of Mobile Advertising*, *2*(1), 76–83.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *The American Psychologist*, *42*, 873. doi:10.1037/h0092053 PMID:2221571

Rossiter, J. R., Silberstein, R. B., Harris, P. G., & Nield, G. (2001). So What? A Rejoinder to the Reply by Crites and Aikman-Eckenrode to Rossiter et al. (2001). *Journal of Advertising Research*, *41*(3), 59–61. Advance online publication. doi:10.2501/JAR-41-3-59-61

Russell, J. A., Lewicka, M., & Niit, T. (1989). A Cross-Cultural Study of a Circumflex Model of Affect. *Journal of Personality and Social Psychology*, *57*(5), 848–856. doi:10.1037/0022-3514.57.5.848

Schmidt, L., Skvortsova, V., Kullen, C., Weber, B., & Plassmann, H. (2017). How context alters value: The brain's valuation and affective regulation system link price cues to experienced taste pleasantness. *Scientific Reports*, *7*(1), 8098. Advance online publication. doi:10.103841598-017-08080-0 PMID:28808246

Shin, D. (2010). The effects of trust, security & privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, *22*(5), 428–438. doi:10.1016/j.intcom.2010.05.001

Singler, E. (2015). *Green Nudge: Changer les comportements pour sauver la planète*. Pearson.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Smarandescu, L., & Shimp, T. A. (2015). Drink Coca-Cola, eat popcorn, and choose powerade: Testing the limits of subliminal persuasion. *Marketing Letters*, *26*(4), 715–726. doi:10.100711002-014-9294-1

Sofi, S. A., Nika, F. A., Shah, M. S., & Zarger, A. S. (2018). Impact of subliminal advertising on consumer buying behaviour: An empirical study on young Indian consumers. *Global Business Review*, *19*(6), 1580–1601. doi:10.1177/0972150918791378

Soler, A. S., Ort, M., Steckel, J., & Nieschullze, J. (2016). *An introduction to data management*. Accessed from: https://www.gfbio.org/documents/10184/22817/Reader_GFBio_BefMate_20160222/1ca43f24-255044b3-a05e-e180c3e544c0

SolomonR. C. (2017). *Emotion*. Retrieved from http://www.britannica.com/topic/emotion

Stoicescu, C. (2016). Big Data, the perfect instrument to study today's consumer behavior. *Database System Journal*, *6*, 28–42.

Stoll, M., Baecke, S., & Kenning, P. (2008). What they see is what they get? An fMRI-study on neural correlates of attractive packaging. *Journal of Consumer Behaviour*, *7*(4-5), 342–359. Advance online publication. doi:10.1002/cb.256

Sultana, N., & Nangunoori, R. (2008). Ethical Acceptability of Neuromarketing: Relevance, Limits and Limitations. *Prabandhan: Indian Journal of Management*, *1*(1), 31. Advance online publication. doi:10.17010//2008/v1i1/64654

Vecchiato, G., Toppi, J., Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Bez, F., & Babiloni, F. (2011). Spectral EEG frontal asymmetries correlate with the experienced pleasantness of video commercial advertisements. *Medical & Biological Engineering & Computing*, *49*(5), 579–583. doi:10.100711517-011-0747-x PMID:21327841

Westbrook, R. A., & Oliver, R. L. (1991). The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction. *The Journal of Consumer Research*, *18*(1), 84–91. doi:10.1086/209243

Wieckowski, A. (2019). When neuromarketing crosses the line. *Harvard Business Review*. Accessed from: https://hbr.org/2019/01/when-neuromarketing-crosses-the-line?ab=at_articlepage_recommendedarticles_bottom1x1

Xu, H., Luo, X., Carroll, J. M., & Rosson, M. B. (2011). The Personalization Privacy Paradox. *Decision Support Systems*, *51*(1), 42–52. doi:10.1016/j.dss.2010.11.017

KEY TERMS AND DEFINITIONS

Artificial Intelligence (AI): AI is a set of algorithms capable of collecting and interpreting data, in order to predict consumer behaviors.

Data Management: The planning, development, implementation, and administration of data systems in order to collect, store, secure, use and archive data.

Eye Tracking: Recording fixation points and eye movements when users look at visual stimuli such as advertising, websites, packaging, or in-store shelves.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Facial Coding: The decoding of specific facial behaviors, based on the facial muscles that produce them, to identify unconscious thoughts and behaviors.

Gestural Coding: The decoding of body attitudes and postures, based on the body muscles that produce them, to identify unconscious thoughts and behaviors.

Neuroethics: The study of the moral values of neuroscientific technologies and the ethics of neuroscience (i.e., the ethical issues raised by neural imaging techniques or cognitive enhancement) (Evers, 2014).

Neuromarketing 2.0: The study of consumers' sensory-motor, cognitive and affective responses to marketing stimuli thanks to AI tools.

Neuroscience: The study of the nervous system, which controls every aspect of the body, from emotion and memory to basic body activities such as movements, breathing and heartbeat.

Pupil Tracking: The recording of the variation of pupil dimensions when consumers look at visual stimuli to identify unconscious perceptions and behaviors.

Vocal Coding: The decoding of the voice, based on characteristics such as tone and timbre, to identify unconscious thoughts and behaviors.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 30
Responsible Machine Learning for Ethical Artificial Intelligence in Business and Industry

Deepak Saxena
https://orcid.org/0000-0002-9331-3799
Birla Institute of Technology and Science, Pilani, India

Markus Lamest
Trinity College Dublin, Ireland

Veena Bansal
Indian Institute of Technology, Kanpur, India

ABSTRACT

Artificial intelligence (AI) systems have become a new reality of modern life. They have become ubiquitous to virtually all socio-economic activities in business and industry. With the extent of AI's influence on our lives, it is an imperative to focus our attention on the ethics of AI. While humans develop their moral and ethical framework via self-awareness and reflection, the current generation of AI lacks these abilities. Drawing from the concept of human-AI hybrid, this chapter offers managerial and developers' action towards responsible machine learning for ethical artificial intelligence. The actions consist of privacy by design, development of explainable AI, identification and removal of inherent biases, and most importantly, using AI as a moral enabler. Application of these action would not only help towards ethical AI; it would also help in supporting moral development of human-AI hybrid.

INTRODUCTION

Breaking away from the realm of science fiction, artificial intelligence (AI) systems have become the reality of modern life. Kaplan and Haenlein (2019) define AI as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks

DOI: 10.4018/978-1-7998-6985-6.ch030

Copyright © 2021, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

through flexible adaptation". AIs in modern life not only provide recommended course of actions to those working in business and industry, but they are also effectively making decisions for us. In our everyday life, they strongly influence what we watch during our leisure time, who we interact more with on social networks, the videos we watch and the music we listen to, what we buy on e-commerce websites and on what rate, what interest rate and credit limits we are offered on our credit cards, what route our taxis take and so on. The list of AI applications is endless. However, AI models are usually opaque, and one often knows very little on the decisions-making algorithm used by AIs (Burrell, 2016; de Saint Laurent, 2018; McQuillan, 2018; Mackenzie, 2019). Since AIs have so much influence on our lives, it is an imperative to focus our attention on the ethics of AI.

At the same time, AIs are built by humans, often using the process of machine learning. Hence, the choices made by the AI developers (who design the AI) and the managers (who decide to implement the recommendations from the AI) are crucial. This paper offers some suggestions towards responsible machine learning building on the idea of human-AI hybrid (Bansal et al., 2019a, 2019b; Dellermann et al., 2019; Jarrahi, 2018; Peeters et al., 2020). Taking a socio-technical systems perspective (Andras et al., 2018), the term human-AI hybrid indicates that each subsystem (AI and the humans) brings a unique skillset to the table and can help in ethical development of the overall system. In so doing, the chapter takes machine learning as a common denominator of AI and introduces a responsible machine learning framework towards ethical AI. For developers and managers, the chapter suggests actionable insights that can be used in industry and business.

The remainder of the chapter is as follows. The next section provides some background in terms of the notion of artificial intelligence, the machine learning process, and the concept of distributed morality in the context of socio-technical systems. Thereafter, the building blocks of responsible machine learning are discussed with the implications they have for the AI developers and managers. The building blocks are aligned with the classical machine learning stages. Based on the discussion, recommendations are presented for the managers and developers. This is followed by noting some limitations and future development, with last section concluding the chapter.

BACKGROUND

The Notion of Artificial Intelligence

While the notion of AI has been popular from the movies like *The Terminator* or *Ex Machina*, there are two crucial limitations related to the popular perception of AI. The first limitation relates to the notion of intelligence. Strictly speaking, the use of the term intelligence in AI does not relate to the entirety of human intelligence but refers mostly to the logical-mathematical intelligence (Brock, 2018; Vetrò et al., 2019). When developers build an AI, they try to automate and optimize the processing that requires data manipulation and logical reasoning. While there have been efforts to design emotionally intelligent AI (Schuller & Schuller, 2018) or AIs with kinaesthetic intelligence (e.g., self-driving cars), the efforts are yet to see the light of the day in wider society. It is with this caution the term 'intelligence' is to be used for AI. Therefore, while AIs are fit to replace humans in the tasks that require an application of logical-mathematical intelligence, human intervention is still needed (Brock, 2018; Ross, 2018; Wilson et al., 2017) due to other forms of intelligence they possess, especially when it comes to ambiguous situations and inductive reasoning. For this reason, this paper draws from the idea of the human-AI hybrid

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Dellermann et al., 2019; Jarrahi, 2018; Peeters et al., 2020) instead of the notion of standalone machine intelligence. Interestingly, this idea is not entirely new but teased upon in the context of expert systems and knowledge engineering (Brock, 2018).

The second limitation relates to the scope of AI. In the popular discourse AI systems are often conceived as a single entity, sometimes termed strong AI (de Saint Laurent, 2018), especially when it comes to issues such as AI ethics. However, what is often ignored is that an AI is built to perform a specific task. It exceeds human capabilities not at an overall level but in the context of information processing for a specific problem (Brock, 2018; de Saint Laurent, 2018). An AI for the bank fraud detection cannot predict the onset of cancer. IBM deep blue may be an excellent chess player; it cannot be used in the autonomous cars to find its way. Therefore, instead of considering AI as an all-encompassing term, one needs a more specific notion to talk about the ethics of AI in business and industry. This common denominator could be machine learning process though which an AI system usually develops its intelligence. It may be noted, however, the machines learning is only an enabler of Artificial Intelligence, and not equivalent to it (de Saint Laurent, 2018). Moreover, while AI ethics as a whole remains a complex subject with no simple answers, machine learning offers some leeway in terms of corrective actions (de Saint Laurent, 2018). The next section briefly explores the machine learning process.

How AI Systems Gain Intelligence

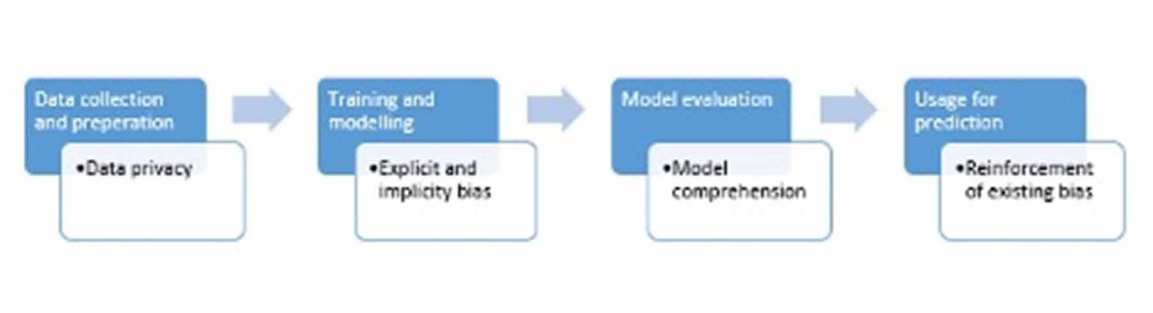
Machine Learning (ML) may be defined as the statistical method that creates models based on training data and continuously improves the model based on the match between the prediction and actual data. In other words, ML programs seek to understand and replicate how certain inputs lead to certain outputs to assist in the task of prediction (de Saint Laurent, 2018). The classical machine learning process is supervised in the sense that the training data is often structured and contains clear inputs and outputs. In contrast, under unsupervised ML, the data is not structured, and the machine looks for patterns in the data without any human intervention. Finally, deep learning is considered as a specific ML method that tries to mimic the functioning of human brain by creating multilayer neural networks, with each layer building on the previous (LeCun et al., 2015; Rusk, 2016). As business examples, while Google DeepMind uses deep learning neural networks, IBM Watson relies on machine learning. While deep learning algorithms are considered more efficient and require more computational power, machine learning models are easy to implement and interpret. In this chapter we focus on classical supervised ML for simplicity. Classical ML may be understood to be consisting of four stages – data collection and preparation, training and modelling, model evaluation, and usage for prediction. Although it is shown in logical sequence, it is useful to remember that the ML process is iterative as AI systems continuously improve the model based on the match between the prediction and actual data. Figure 1 outlines each ML stage and notes potential issues from an ethical perspective.

Data Collection and Preparation

Existing datasets are the starting point for ML algorithms to train. It is argued that the ML race between companies is essentially a race about data (Beck & Libert, 2019). The quality of the AI depends upon the amount and the quality of data used to train the AI. ML provides an opportunity to learn from traditional data such as browsing or purchase as well as from non-traditional sources such as big data (Zhou et al., 2017) to support business decisions. This is why companies prefer to collect user data from as many

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Machine learning stages and potential issues



sources as possible. At the same time, it is observed that big data results in an information overload for managers (Saxena & Lamest, 2018) due to its volume and velocity. ML provides a more efficient solution to the problem of making sense of the wide tapestry of data residing within and outside the company. However, the database design process almost exclusively focuses upon database and query optimization, virtually disregarding issues such as privacy or trust. With the technology infrastructure becoming all pervasive and ever cheaper, there are obvious tensions between convenience versus privacy (Weinberg et al., 2015) and customer insights versus privacy intrusion (Plangger & Watson, 2015). Especially with the advent of ubiquitous and non-intrusive technologies such as Internet of Things (IoT), not only customers need to be aware of the privacy invasion, but organizations also need to be mindful of the extent of data collection and its wider repercussions (Santanen, 2019) to avoid reputation loss and legal challenges.

Training and Modelling

This is this stage when AI systems use complex statistical methods for pattern recognition and to make sense of the training data (de Saint Laurent, 2018). The result of this process is often a model to explain the specific patterns and to make predictions. While there are a variety of approaches to optimize the learning process (e.g., see Qiu et al., 2017), it is useful to remember that an AI model draws from the training data fed to it and essentially reinforces the existing patterns from which training data is generated (de Saint Laurent, 2018). Consequently, while the model might be statistically optimal with respect to the training data fed to it, it might still have inherent biases (Daugherty et al., 2019; Miller, 2020; Vetrò et al., 2019; Žliobaitė, 2015) due to disproportionate datasets and/or collinearity. Training on the past data runs the risk of introducing certain bias in the model (Ajana, 2015; de Saint Laurent, 2018; Turner Lee, 2018). For instance, few years back a leading business needed to shut down its recruitment AI because it seemed to disadvantage female candidates. Since women are traditionally under-represented in the technology sector, AI systems usually inherit gender bias (Leavy, 2018). Similarly, Ajana (2015) reports that the use of big data and machine learning in border operations runs the risk of creating undue restrictions on certain 'high-risk' identities and privileging certain identities. More recently, UK authorities had to revise the school grades predicted by an AI since it downgraded some grades assigned by the teachers, based on the school profile (Evgeniou et al., 2020) even though the teachers were confident in their grade assignment. These examples clearly illustrate that the importance of identification and removal of biases at the training and modelling stage for non-discriminative use of data in ML (Žliobaitė, 2015).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Model Evaluation

The next step in the ML process is to evaluate the efficacy of the model. This is usually evaluated with reference to the fit between the predicted outcomes and the actual outcomes. However, statistical fit is only one aspect of model evaluation. A key managerial challenge is to make sense of the models themselves. Since there might be thousands of variables and complex algebraic expressions at play, sometimes managers step in the dark (de Saint Laurent, 2018; McQuillan, 2018; Wilson et al, 2017) regarding how the model actually works. For instance, it was reported in 2015 that a popular photo application was tagging people of color as Gorilla. After unsuccessfully trying to address the issue, the eventual solution used by the developers was to remove the term gorilla (and neighboring terms – chimpanzee, monkey etc.) from its lexicon. This implies that sometimes even the developers may struggle in understanding the decision-model used by the system. This issue becomes extremely complex when there are competing models and algorithms to choose from. Thus, choosing a model in a specific context itself becomes an ethical exercise (Webb et al., 2019) for decision makers. The issue becomes even more complex with neural network based deep learning models that are rarely comprehensible to humans (Dickson, 2018). Consequently, model selection often becomes a matter of experience and subjective impression of the researcher (de Saint Laurent, 2018).

Usage for Prediction

Just as computers' main task is to execute efficient calculations, the task of an AI is to facilitate predictions (Agrawal et al., 2019). However, unlike traditional computer programs that are expected to repeat the same steps that are designed by their programmers, an AI keeps modifying its logic based on the relative success/failure of its prediction. Irrespective of the model and logic, however, one central limitation is that although the past data may be used to predict the future; it should not determine the future (Ajana, 2015). As Kerr and Earle (2013) argue, by reinforcing past bias, this approach restricts the future possibilities for those who are subject to this algorithmic governmentality (Harkens, 2018), for instance limiting jobs or immigration prospects. Thus, we need human intervention at this stage to make a call on whether the recommendations made by AI are fair and safe (Bansal et al., 2019a; "Perspectives on Issues in AI Governance", 2019).

Throughout the machine learning process, managers and developers are involved in different capacities. In this context, the next section discusses the notion of distributed morality and human-AI hybrid.

Distributed Morality and the Human-AI Hybrid

Although popular literature may wish to assign agency to an AI program, there is an increasing support for the notion of distributed morality (Floridi, 2013; de Saint Laurent, 2018; Magnani & Bardone, 2008) for complex information systems such as an AI. Towards the goal of ethical and trustworthy AI, national and transnational AI strategies and guidelines seek to provide a regulatory framework and set of principles (Dafoe, 2018; Lauterbach, 2019) for AI governance. Recent guidelines on trustworthy AI from the European Commission (2019) and the Algorithmic Accountability Bill presented in the US Congress (2019) are two such examples. Similarly, Ethically Aligned Design from the Institute of Electrical and Electronics Engineers (IEEE, 2019) provides a broad set of principles for the design and governance

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of AI. While such principles provide a broad framework for ethical AI governance, AI developers and business users still need to take decisions in their specific business context.

Distributed morality implies that ethics of a complex system are determined by the actions taken by various actors involved in the development, implementation, and use of the system. In this respect, the first burden of morality falls on AI developers. Technology developers usually build a technology with a defined functionality in mind, usually dictated by the organizational needs. However, this does not exonerate them from their ethical responsibility. Since the affordances designed in the technological artefact support or constrain certain actions, AI developers end up doing ethics by other means (Verbeek, 2006, 2011). The second burden of morality falls on the business users. An AI is designed towards specific business goals that need to be framed within a broader societal context. Hence, not only the goal determination is an inherent ethical process, but also is the choice of whether to act on AI's recommendations. Even saying 'yes' to AI recommendations constitutes an ethical decision on part of the business users (de Saint Laurent, 2018).

Thus, as opposed to the notion of an AI working in isolation, it is the human-AI hybrid system (Andras et al., 2018; Bansal et al., 2019a; Dellermann et al., 2019; Jarrahi, 2018; Peeters et al., 2020; "Perspectives on Issues in AI Governance", 2019) one needs to focus upon. While the AIs have an immense capacity for logical-mathematical processing (Vetrò et al., 2019), they are not conceived as possessing awareness and intentionality (de Saint Laurent, 2018). In contrast, reflection and moral self-awareness are arguably a crucial part of moral development (Friedland & Cole, 2019) in humans and collectives. At the current stage of development, AI systems do not possess a capability for self-awareness or reflection. Thus, following-on from the idea of distributed morality and human-AI hybrid, the key for AI ethics lies in humans facilitating the ML process to incorporate ingredients of reflection and moral self-awareness to turn the AI into socially responsible agents (Vetrò et al., 2019). The next section discusses those actions required for responsible machine learning.

TOWARDS RESPONSIBLE MACHINE LEARNING

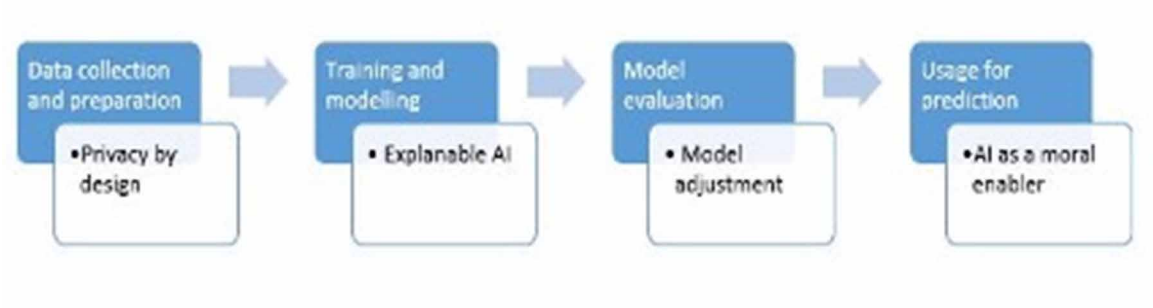
Aligned with the notion of human-AI hybrid and distributed morality, Figure 2 shows a responsible ML model with corrective actions. In a limited manner, it draws from knowledge discovery in databases (KDD) approach that involves data preprocessing, data cleaning, incorporation of prior knowledge, and appropriate interpretation/evaluation when extracting knowledge from available datasets (Custers, 2013; Fayyad et al., 1996). In a limited manner, it tries to bring classical ML closer to the KDD process in which knowledge must be more understandable and useful than being optimal and precise (Kodratoff, 1999).

Privacy by Design

The first and the most basic step for a responsible AI is putting privacy and fairness at the core of its design. At the preprocessing stage, training data may be analyzed for inherent biases for the purpose of removing those before feeding into ML training stage (Žliobaitė, 2015). A more effective measure, however, may be introduced at the data collection stage in the form of privacy by design. The recent enactment of the General Data Protection Regulation (GDPR) in the European Union facilitates and promotes user privacy and puts a legal binding on the organizations. For managers and employees, this means that they should be aware of the data protection and privacy regulations, and do not put demands

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Towards responsible machine learning



on the AI developers that violate these principles. Especially with the advent of big data and IoT, not only customers need to be aware of the privacy invasion; businesses also need to be mindful of the extent of data collection and its wider repercussions.

The notion of Privacy by Design (Bednar, 2019; Cavoukian, 2012; Wahlstrom et al., 2020) may be employed during the data preparation stage, incorporating privacy affordances in the planning and the design of the system itself (Matwin, 2013). For example, different privacy levels may be used to reflect the privacy sensitivity of the data. This would help in data minimization (Schaar, 2010) i.e., no or minimal collection/processing of individually identifiable information but efficient utilization at the aggregate level. In this regard, the Algorithmic Accountability bill presented in the US Congress (2019) mandates businesses to provide a detailed description of the AI, its algorithmic design, training data used, and the purpose of the system. It also requires them to prepare and report an impact assessment study on the accuracy, fairness/biases, privacy, and security of the system.

Explainable AI

Privacy is only the first stage of an ethical ML process. To ensure an ethical human-AI hybrid, business users should have an adequate understanding of the actual working, purpose, and the limitations of the AI (de Saint Laurent, 2019; Martinho-Truswell, 2018). In this regard, the recent European Commission (2019) guidelines on trustworthy AI suggest appointing a responsible person, or in some cases even an entire panel or board for this purpose. The guidelines suggest that in case an AI system has a significant impact on people's lives, even a layperson should be able to understand both the causality of the algorithmic decision-making process and how this is implemented by the respective organization. In this regard, Wilson et al (2017) identify a new job category of 'explainers' who can bridge the gap between the AI and managerial understanding.

However, it is easier said than done (Weld & Bansal, 2019). ML models are so complex that the managers (sometimes even the developers) may not have enough understanding of the underlying logic. To have full trust in the system, a manager needs to know on what basis and how the AI comes up with a specific recommendation (Ross, 2018). To support this requirement, we need machine learning models that can be broadly explained to humans, often called an explainable AI (Dickson, 2018; Hagras, 2018; Rai, 2020) or in short XAI. An XAI may take diverse approaches. It may involve explaining the building blocks of the ML model, explaining the reasoning, or developing a text-based post-hoc explanation (Arrieta et al., 2020). There arises a question of the trade-off between the explicability of the AI and its

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

efficacy (Brock, 2018; London, 2019; "Perspectives on Issues in AI Governance", 2019) as too much focus on explanation might defeat the purpose of AI creation for optimal and quick decision-making. At the same time, however, Bansal et al. (2019b) demonstrate that (standalone) AI performance improvements may actually hurt the performance of human-AI team.

Model Adjustment

Despite the efforts towards explainable AI, there is no denying that certain explicit or implicit biases (Miller, 2020; Turner Lee, 2018) will be present in the model. In this regard, Žliobaitė (2015) suggests model postprocessing at this stage in terms of changing certain parameters. Moreover, rather than treating these biases as an unwanted outcome, developers and managers may use the AI as a detection mechanism to surface biases ("Perspectives on Issues in AI Governance", 2019). Once an AI detects a specific bias, managers and developers can take steps to remove specific biases, or even to introduce affirmative biases. For instance, for a recruitment AI, one may remove variables such as house address (that may bias an AI towards candidates residing in affluent areas) and/or introduce positive biases for underrepresented populations (e.g., women and third gender, people of color). At the managerial level, this would require discussing and determining which existing biases need to be removed and which positive biases or affirmative actions need to be introduced in the model (Verwer & Calders, 2013). For AI developers, this would mean allowing for changing weights on certain decision criteria or conducting a 'what-if' analysis. However, this should be done on a continuous basis since Calders and Žliobaitė (2013) demonstrate that such biases may creep-in in the model even when relevant variables are removed from the dataset. In this regard, Wilson et al (2017) identify a new job category of 'sustainers' who will ensure that an AI system performs as intended and unintended outcomes are dealt with utmost urgency. Thus, frequent debates (de Saint Laurent, 2018) on AI's decisions and ML models is required for such model adjustments.

Using AI as a Moral Enabler

A crucial implication of human-AI hybrids is that they evolve together. While an AI helps humans in sifting through complex datasets, a human may help AI in working with uncertainty and equivocality of business decisions (Wilson et al., 2017). Moreover, we argue that AI could be used as a moral enabler (Floridi, 2013) for humans. This would involve introducing opportunity for reflection and self-awareness. AI developers may introduce AI friction filter via visceral notices (Friedland, 2019) to trigger the reflective mind. For instance, the managers and developers may be shown the visualization of the outcome variable with key determining variables ("Perspectives on Issues in AI Governance", 2019). The users may then use the opportunity to reflect on whether they wish to go with a specific model or wish to further improve it. Another approach might be informing the end-user about a different set of action than that recommended based on the past behavior. IEEE (2019) recommends that such comprehensive information should be provided to recognize different kind of algorithmic interventions on an opt-in basis. The AI may prompt the users if they are accepting all automated decisions without spending enough time reflecting on the options available to them. As Lara and Deckers (2019) suggest, by constantly asking reflective questions, an AI may act as a Socratic assistant for moral enhancement of the human-AI hybrid. It may also work the other way round when humans assist an AI in improving its trustworthiness. For example, Zhang et al (2019) have proposed a hybrid system called *CrowdLearn*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

that complements the machine intelligence with the intelligence of the crowd to detect anomalies and improve opaque AI algorithms. Thus, humans and AI can help each other towards moral development of the hybrid system (Bansal et al., 2019a).

SOLUTIONS AND RECOMMENDATIONS

Based on the discussion above, Table 1 outlines the actions recommended for responsible machine learning in business and industry. However, two caveats must be noted here. First, each ML stage is iterative in nature to support learning and to facilitate reflection. Second, although the actions for AI developers and managers are shown in separate rows, in practice they need to collaborate and support each other on a continuous basis. In a sense, the developers and the managers form the feedback loop for each ML stage.

Table 1. Recommended actions for responsible machine learning

Recommended Actions	Privacy by Design	Explainable AI	Model Adjustment	AI as a Moral Enabler
Action at the level of managers	• Determine the data and privacy protection policies. • Understand and explain the use of data to the employees.	• Those using the results from AI must roughly understand what goes inside the black box. • Determine which models/ constructs are to be used.	• Determine which existing biases are unwanted, and which affirmative biases are to be introduced.	• Reflect on the decision-making process and its impact.
Actions at the level of developers	• Design systems in a way that allows for further protection of privacy.	• Facilitate explainable AI.	• Make existing biases visible through explainable AI. • If existing biases are unwanted, exclude certain variables from the decision criteria. • Introduce affirmative biases in the system by changing the decision criteria.	• Introduce opportunities for reflection.

FUTURE RESEARCH DIRECTIONS

Certain limitations of this work, that inform the future research directions, may be noted. One limitation of the chapter is that its recommendations are mainly applicable to classical machine learning approaches with linear algorithmic assumptions. Deep learning approaches that rely on hierarchical neural network analysis, are often difficult to explain and interpret. Thus, a dedicated approach to responsible deep learning (such as the one illustrated in Angelov & Soares, 2020) is warranted. Having noted that, the framework described in this chapter may be used across a diverse range of ML applications and need not be confined to AI development. Second, despite being actionable, the recommendations remain conceptual and need to be examined empirically. These recommendations need to be informed by practice, in the form of case study of AI development or a survey of experts in the future work. Finally, it needs to be acknowledged that the discourse on AI ethics is ever evolving and requires continuous interaction and debate among the developers, regulators, business and civil society.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

AIs are increasingly becoming a reality in modern life. With all pervasiveness of AI systems that can collect information on our minute actions and process complex information sets ranging from IoT to Big Data, there is a need for a framework that attends to the ethical aspect of an AI. Self-awareness and reflection help in the moral development of humans. Despite their immense capacity for logical-mathematical processing, the current generation of AIs lack the ability to reflect and can rarely be called self-aware. Focusing on the machine learning, the process through which an AI develops its intelligence, we offer a set of actions towards responsible machine learning, outlining managerial and developers' actions at each ML stage. Responsible machine learning may include privacy by design, development of explainable AI, identification and removal of inherent biases, and most importantly, use of AI as a moral enabler. We trust that these actions would not only help in the development of ethical AI systems, but also would also help in supporting further moral development of combined human-AI hybrids.

REFERENCES

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy*, *47*, 1–6. doi:10.1016/j.infoecopol.2019.05.001

Ajana, B. (2015). Augmented borders: Big Data and the ethics of immigration control. *Journal of information*. *Communication and Ethics in Society*, *13*(1), 58–78. doi:10.1108/JICES-01-2014-0005

Andras, P., Esterle, L., Guckert, M., Han, T. A., Lewis, P. R., Milanovic, K., Payne, T., Perret, C., Pitt, J., Powers, S. T., Urquhart, N., & Wells, S. (2018). Trusting intelligent machines: Deepening trust within socio-technical systems. *IEEE Technology and Society Magazine*, *37*(4), 76–83. doi:10.1109/MTS.2018.2876107

Angelov, P., & Soares, E. (2020). Towards explainable deep neural networks (xDNN). *Neural Networks*, *130*, 185–194. doi:10.1016/j.neunet.2020.07.010 PMID:32682084

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Chatila, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. doi:10.1016/j.inffus.2019.12.012

Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019a). Beyond accuracy: The role of mental models in human-ai team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (*Vol. 7*, No. 1, pp. 2-11). Academic Press.

Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., & Horvitz, E. (2019, July). Updates in human-ai teams: Understanding and addressing the performance/compatibility tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 2429–2437. doi:10.1609/aaai.v33i01.33012429

Beck, M., & Libert, B. (2019). The machine learning race is really a data race. *MIT Sloan Management Review*. Available at https://sloanreview.mit.edu/article/the-machine-learningrace-is-really-a-data-race/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bednar, K., Spiekermann, S., & Langheinrich, M. (2019). Engineering Privacy by Design: Are engineers ready to live up to the challenge? *The Information Society*, *35*(3), 122–142. doi:10.1080/01972243.2019.1583296

Brock, D. C. (2018). Learning from Artificial Intelligence's Previous Awakenings: The History of Expert Systems. *AI Magazine*, *39*(3), 3–15. doi:10.1609/aimag.v39i3.2809

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 2053951715622512. doi:10.1177/2053951715622512

Calders, T., & Žliobaitė, I. (2013). Why unbiased computational processes can lead to discriminative decision procedures. In *Discrimination and privacy in the information society* (pp. 43–57). Springer. doi:10.1007/978-3-642-30487-3_3

Cavoukian, A. (2012). Privacy by design: origins, meaning, and prospects for assuring privacy and trust in the information era. In G. O. M. Yee (Ed.), *Privacy protection measures and technologies in business organizations: aspects and standards* (pp. 170–208). IGI Global. doi:10.4018/978-1-61350-501-4.ch007

Custers, B. (2013). Data dilemmas in the information society: Introduction and overview. In *Discrimination and Privacy in the Information Society* (pp. 3–26). Springer. doi:10.1007/978-3-642-30487-3_1

Dafoe, A. (2018). AI governance: a research agenda. Governance of AI Program, Future of Humanity Institute, University of Oxford.

Daugherty, P. R., Wilson, H. J., & Chowdhury, R. (2019). Using artificial intelligence to promote diversity. *MIT Sloan Management Review*, *60*(2), 1.

De Saint Laurent, C. (2018). In defence of machine learning: Debunking the myths of artificial intelligence. *Europe's Journal of Psychology*, *14*(4), 734–747. doi:10.5964/ejop.v14i4.1823 PMID:30555582

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643. doi:10.100712599-019-00595-2

Dickson, B. (2018). Explainable AI: Learning how AI makes decisions. *PC Mag*. Available at https://uk.pcmag.com/news/118591/learning-how-ai-makes-decisions

European Commission. (2019), *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Evgeniou, T., Hardoon, D. R., & Ovchinnikov, A. (2020). What Happens When AI is Used to Set Grades? *Harvard Business Review*. https://hbr.org/2020/08/what-happens-when-ai-is-used-to-set-grades

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, *39*(11), 27–34. doi:10.1145/240455.240464

Floridi, L. (2013). Distributed morality in an information society. *Science and Engineering Ethics*, *19*(3), 727–743. doi:10.100711948-012-9413-4 PMID:23197312

Friedland, J. (2019). AI Can Help Us Live More Deliberately. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/ai-can-help-us-live-more-deliberately

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Friedland, J., & Cole, B. M. (2019). From homo-economicus to homo-virtus: A system-theoretic model for raising moral self-awareness. *Journal of Business Ethics*, *155*(1), 191–205. doi:10.100710551-017-3494-6

Hagras, H. (2018). Toward human-understandable, explainable AI. *Computer*, *51*(9), 28–36. doi:10.1109/MC.2018.3620965

Harkens, A. (2018). The ghost in the legal machine: Algorithmic governmentality, economy, and the practice of law. *Journal of Information Communication and Ethics in Society*, *16*(1), 16–31. doi:10.1108/JICES-09-2016-0038

IEEE. (2019). *Ethically Aligned Design*. Institute of Electrical and Electronic Engineers. https://ethicsinaction.ieee.org/

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. doi:10.1016/j.bushor.2018.03.007

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. doi:10.1016/j.bushor.2018.08.004

Kerr, I., & Earle, J. (2013). Prediction, preemption, presumption: How big data threatens big picture privacy. *Stanford Law Review Online*, *66*, 65–72.

Kodratoff, Y. (1999, July). Comparing machine learning and knowledge discovery in databases: An application to knowledge discovery in texts. In *Advanced Course on Artificial Intelligence* (pp. 1–21). Springer.

Lara, F., & Deckers, J. (2019). Artificial Intelligence as a Socratic Assistant for Moral Enhancement. *Neuroethics*, 1–13. doi:10.100712152-019-09401-y

Lauterbach, A. (2019). Artificial intelligence and policy: Quo vadis? *Digital Policy*. *Regulation & Governance*, *21*(3), 238–263. doi:10.1108/DPRG-09-2018-0054

Leavy, S. (2018, May). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In *Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14-16). 10.1145/3195570.3195580

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

London, A. J. (2019). Artificial intelligence and black-box medical decisions: Accuracy versus explainability. *The Hastings Center Report*, *49*(1), 15–21. doi:10.1002/hast.973 PMID:30790315

Mackenzie, A. (2019). From API to AI: Platforms and their opacities. *Information Communication and Society*, *22*(13), 1989–2006. doi:10.1080/1369118X.2018.1476569

Magnani, L., & Bardone, E. (2008). Distributed morality: Externalizing ethical knowledge in technological artifacts. *Foundations of Science*, *13*(1), 99–108. doi:10.100710699-007-9116-5

Martinho-Truswell, E. (2018), 3 Questions About AI That Nontechnical Employees Should Be Able to Answer. *Harvard Business Review*. https://hbr.org/2018/08/3-questionsabout-ai-that-nontechnicalemployees-should-be-able-to-answer

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Matwin, S. (2013). Privacy-preserving data mining techniques: survey and challenges. In *Discrimination and Privacy in the Information Society* (pp. 209–221). Springer. doi:10.1007/978-3-642-30487-3_11

McQuillan, D. (2018). People's councils for ethical machine learning. *Social Media + Society*, *4*(2), 2056305118768303. doi:10.1177/2056305118768303

Miller, K. (2020). A Matter of Perspective: Discrimination, Bias, and Inequality in AI. In M. Jackson & M. Shelly (Eds.), Legal Regulations, Implications, and Issues Surrounding Digital Data (pp. 182-202). IGI Global.

Peeters, M. M., van Diggelen, J., Van Den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2020). Hybrid collective intelligence in a human–AI society. *AI & Society*, 1–22.

Perspectives on Issues in AI Governance. (2019). *Google*. https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf.

Plangger, K., & Watson, R. T. (2015). Balancing customer privacy, secrets, and surveillance: Insights and management. *Business Horizons*, *58*(6), 625–633. doi:10.1016/j.bushor.2015.06.006

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 67. doi:10.118613634-016-0355-x

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. doi:10.100711747-019-00710-5

Ross, J. (2018). The fundamental flaw in AI implementation. *MIT Sloan Management Review*, *59*(2), 10–11.

Rusk, N. (2016). Deep learning. *Nature Methods*, *13*(1), 35–35. doi:10.1038/nmeth.3707 PMID:27110626

Santanen, E. (2019). The value of protecting privacy. *Business Horizons*, *62*(1), 5–14. doi:10.1016/j.bushor.2018.04.004

Saxena, D., & Lamest, M. (2018). Information overload and coping strategies in the big data context: Evidence from the hospitality sector. *Journal of Information Science*, *44*(3), 287–297. doi:10.1177/0165551517693712

Schaar, P. (2010). Privacy by design. *Identity in the Information Society*, *3*(2), 267–274. doi:10.100712394-010-0055-x

Schuller, D., & Schuller, B. W. (2018). The age of artificial emotional intelligence. *Computer*, *51*(9), 38–46. doi:10.1109/MC.2018.3620963

Turner Lee, N. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information. Communication and Ethics in Society*, *16*(3), 252–260. doi:10.1108/JICES-06-2018-0056

US Congress. (2019). *HR 2231 – Algorithmic Accountability Act of 2019*. https://www.congress.gov/bill/116th-congress/house-bill/2231/text

Verbeek, P. P. (2006). Materializing morality: Design ethics and technological mediation. *Science, Technology & Human Values*, *31*(3), 361–380. doi:10.1177/0162243905285847

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press. doi:10.7208/chicago/9780226852904.001.0001

Verwer, S., & Calders, T. (2013). Introducing positive discrimination in predictive models. In *Discrimination and privacy in the information society* (pp. 255–270). Springer. doi:10.1007/978-3-642-30487-3_14

Vetrò, A., Santangelo, A., Beretta, E., & De Martin, J. C. (2019). AI: From rational agents to socially responsible agents. *Digital Policy*. *Regulation & Governance*, *21*(3), 291–304. Advance online publication. doi:10.1108/DPRG-08-2018-0049

Wahlstrom, K., Ul-haq, A., & Burmeister, O. (2020). Privacy by design. *AJIS. Australasian Journal of Information Systems*, *24*, 1–9. doi:10.3127/ajis.v24i0.2801

Webb, H., Patel, M., Rovatsos, M., Davoust, A., Ceppi, S., Koene, A., Dowthwaite, L., Portill, V., Jirotka, M., & Cano, M. (2019). It would be pretty immoral to choose a random algorithm: Opening up algorithmic interpretability and transparency. *Journal of Information Communication and Ethics in Society*, *17*(2), 210–228. doi:10.1108/JICES-11-2018-0092

Weinberg, B. D., Milne, G. R., Andonova, Y. G., & Hajjat, F. M. (2015). Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons*, *58*(6), 615–624. doi:10.1016/j.bushor.2015.06.005

Weld, D. S., & Bansal, G. (2019). The challenge of crafting intelligible intelligence. *Communications of the ACM*, *62*(6), 70–79. doi:10.1145/3282486

Wilson, H. J., Daugherty, P., & Bianzino, N. (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review*, *58*(4), 14.

Zhang, D., Zhang, Y., Li, Q., Plummer, T., & Wang, D. (2019, July). Crowdlearn: A crowd-ai hybrid system for deep learning-based damage assessment applications. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)* (pp. 1221-1232). IEEE. 10.1109/ICDCS.2019.00123

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350–361. doi:10.1016/j.neucom.2017.01.026

Žliobaitė, I. (2015). *A survey on measuring indirect discrimination in machine learning.* arXiv preprint arXiv:1511.00148.

ADDITIONAL READING

Boddington, P. (2017). *Towards a Code of Ethics for Artificial Intelligence*. Springer. doi:10.1007/978-3-319-60648-4

Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. The Cambridge handbook of artificial intelligence, 1, 316-334.

Custers, B. H. M., Calders, T. G. K., Schermer, B. W., & Zarsky, T. Z. (2013). *Discrimination and privacy in the information society: data mining and profiling in large databases*. Springer. doi:10.1007/978-3-642-30487-3

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Flasiński, M. (2016). *Introduction to Artificial Intelligence*. Springer. doi:10.1007/978-3-319-40022-8

Floridi, L. (2013). *The Ethics of Information*. Oxford University Press. doi:10.1093/acprof:oso/9780199641321.001.0001

Jackson, P. C. (2019). *Introduction to Artificial Intelligence*. Courier Dover Publications.

Molnar, C. (2021). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-ml-book/

Sheikh, S. (2021). *Understanding the Role of Artificial Intelligence and Its Future Social Impact*. IGI Global. doi:10.4018/978-1-7998-4607-9

Thompson, S. J. (2021). *Machine Law, Ethics, and Morality in the Age of Artificial Intelligence*. IGI Global. doi:10.4018/978-1-7998-4894-3

Vassileva, B., & Zwilling, M. (2021). *Responsible AI and Ethical Issues for Businesses and Governments*. IGI Global. doi:10.4018/978-1-7998-4285-9

White, J., & Searle, R. (2015). *Rethinking Machine Ethics in the Age of Ubiquitous Technology*. IGI Global. doi:10.4018/978-1-4666-8592-5

KEY TERMS AND DEFINITIONS

Artificial Intelligence: A system that develops a prediction model based on the training data, predicts future trends based on current data, and recommends a future course of action.

Deep Learning: A subset of machine learning in which an AI prepares a multilayer neural network of algorithms to build and constantly refine a prediction model based on available data (may be structured or unstructured).

Explainable AI: An AI for which the underlying model may be explained using the constructs understandable to humans. The explanation may be incorporated during the learning process or may be introduced once the model is built.

Human-AI Hybrid: A concept that views humans and AI as subsystems of a socio-technical system, with the implication that action of either have an influence on the other and on the overall system.

Machine Learning: The learning process through which an AI analyses the training dataset (usually structured), builds, and refines a prediction model. Machine learning may be supervised or unsupervised.

Privacy by Design: A concept that puts privacy at the center of the system design process. At each step of the process, privacy-preserving techniques are introduced to support for data protection.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Compilation of References

Aaker, D. A. (1996). *Building Strong Brands*. Simon & Schuster.

Aaker, D. A., Kumar, V., & Day, G. S. (2008). *Marketing research* (9th ed.). John Wiley & Sons.

Abad-Segura, E., González-Zamar, M. D., de la Rosa, A. L., & Gallardo-Pérez, J. (2020). Management of the digital economy in higher education: Trends and future perspectives [Gestión de la economía digital en la educación superior: tendencias y perspectivas futuras]. *Campus Virtuales*, *9*(1), 57–68.

Abbasi-Sureshjani, S., Raumanns, R., Michels, B. E. J., Schouten, G., & Cheplygina, V. (2020). Risk of training diagnostic algorithms on data with demographic bias. *Lecture Notes in Computer Science*, *12446*, 183–192. doi:10.1007/978-3-030-61166-8_20

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11), 1–41. doi:10.1016/j.heliyon.2018.e00938 PMID:30519653

Abràmoff, M. D., Tobey, D., & Char, D. S. (2020). Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. *American Journal of Ophthalmology*, *214*, 134–142. doi:10.1016/j.ajo.2020.02.022 PMID:32171769

Abratt, R., & Motlana, P. (2002). Managing co-branding strategies: Global brands into local markets. *Business Horizons*, *45*(5), 43–50. doi:10.1016/S0007-6813(02)00242-2

Academic Acelerator. (2021). *Academic Accelerator: Accelerate your scientific research*. Retrieved from https://academic-accelerator.com

Accenture. (2006). *E-government leadership: Building the trust*. Accenture.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access: Practical Innovations, Open Solutions*, *6*, 52138–52160. doi:10.1109/ACCESS.2018.2870052

Adhikari, A., & Bhattacharya, S. (2016). Appraisal of literature on customer experience in tourism sector: Review and framework. *Current Issues in Tourism*, *19*(4), 296–321. doi:10.1080/13683500.2015.1082538

Adobe. (2020). *Digital Trends Study: 2020 Is The Year Of CX-Centric Business Transformation*. Retrieved January 15, 2021, from https://cmo.adobe.com/articles/2020/2/digital-trends-study--2020-is-the-year-of-cx-centric-business-tr.html#gs.imxcnu

Affisco John, F., & Soliman Khalid, S. (2006). E-government: A strategic operations management framework for service delivery. *Business Process Management Journal*, *12*(1), 13–21. doi:10.1108/14637150610643724

Agbaba, R. (2020). Maritime Challenges in Crisis Times. *Annals of Maritime Studies / Pomorski Zbornik, 59*(1), 51–60. . doi:10.18048/2020.59.03

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *The Journal of Consumer Research*, *34*(4), 468–479. doi:10.1086/518544

Aghdaie, N., Kolen, J., Mattar, M., Sardari, M., Xue, S., & Atif-Uz Zaman, K. (2018). *Multiplayer Video Game Matchmaking Optimization*. United States Patent 9993735.

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy*, *47*, 1–6. doi:10.1016/j.infoecopol.2019.05.001

Agyeman, J. K., Ameyaw, B., Li, Y., Appiah-Kubi, J., Annan, A., Oppong, A., & Twumasi, M. A. (2020). Modeling the long-run drivers of total renewable energy consumption: Evidence from top five heavily polluted countries. *Journal of Cleaner Production*, *277*, 123292. doi:10.1016/j.jclepro.2020.123292

Ahmed, A.-O., & Hussein, A.-O. (2006). E-government readiness assessment model. *Journal of Computational Science*, *2*(11), 841–845. Advance online publication. doi:10.3844/jcssp.2006.841.845

Aiello, G., Giallanza, A., & Mascarella, G. (2020). Towards Shipping 4.0. A preliminary gap analysis. *Procedia Manufacturing*, *42*, 24–29. doi:10.1016/j.promfg.2020.02.019

Ainouz, S. A., & Ben Ahmed, M. (2020). *A Smart Chatbot Architecture based on NLP and Machine Learning for Health Care Assistance*. . doi:10.1145/3386723.3387897

Ajana, B. (2015). Augmented borders: Big Data and the ethics of immigration control. *Journal of information. Communication and Ethics in Society*, *13*(1), 58–78. doi:10.1108/JICES-01-2014-0005

Aji, B. M., & Larner, A. J. (2017). Screening for dementia: Single yes/no question or Likert scale? *Clinical Medicine*, *17*(1), 93–94. doi:10.7861/clinmedicine.17-1-93 PMID:28148591

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. Advance online publication. doi:10.1016/0749-5978(91)90020-T

Akerlof, G. A., & Shiller, R. J. (2015). *The economics of manipulation. How we fall like dupes into the plots of the market*. Planeta Colombiana S.A.

Akhter, S. H. (2014). Privacy concern and online transactions: The impact of internet self-efficacy and internet involvement. *Journal of Consumer Marketing*, *31*(2), 118–125. doi:10.1108/JCM-06-2013-0606

Akkiraju, R., Sinha, V., Xu, A., Mahmud, J., Gundecha, P., Liu, Z., Liu, X., & Schumacher, J. (2020). *Characterizing machine learning processes: A maturity framework*. Paper presented at the International Conference on Business Process Management. 10.1007/978-3-030-58666-9_2

Al-adawi, Z., Cardiff, Y., & Pallister, J. (2005). *Conceptual Model of Citizen Adoption of E-Government.* Paper presented at the Second International Conference on Innovations in Information Technology (IIT'05). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.2165&rep=rep1&type=pdf

Alavi, A., Sayareh, J., Fei, J., & Nguyen, H.-O. (2018). Port Logistics Integration: Challenges and Approaches. *International Journal of Supply Chain Management, 7*.

Algorithmia. (2020). *The roadmap to machine learning maturity*. Retrieved from https://info.algorithmia.com/hubfs/2018/Whitepapers/The_Roadmap_to_Machine_Learning_Maturity_Final.pdf

Ali, F., Kim, W. G., Li, J., & Jeon, H. M. (2018). Make it delightful: Customers' experience, satisfaction and loyalty in Malaysian theme parks. *Journal of Destination Marketing & Management*, *7*, 1–11. doi:10.1016/j.jdmm.2016.05.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities (London, England)*, *89*, 80–91. doi:10.1016/j.cities.2019.01.032

Allan Turing: Life and Legacy of a Great Thinker. (n.d.). https://link.springer.com/chapter/10.1007/978-3-662-05642-4_13

Allardice, L. (2021). *Interview: Kazin Ishiguro: AI, gene-editing, big data… I worry we are not in control of these things anymore*. The Guardian. https://www.theguardian.com/books/2021/feb/20/kazuo-ishiguro-klara-and-the-sun-interview?fbclid=IwAR0ExKc5Jda-6TEWM-l-zJzBGYmFgpw0N-DWVrlzPfIl6Ehw3Dh1WPIaRMI

Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality Progress*, *40*(7), 64–65.

Alliance for Qualification. (2019). *AI and Software Testing Foundation Syllabus*. Retrieved from https://www.alliance-4qualification.info/a4q-ai-and-software-testing

Alonso-Cortés, F. (2020*). ¿Hacia dónde vamos en la nueva era del cliente?* Retrieved December 16, 2020, from https://www.marketingdirecto.com/especiales/the-future-of-advertising-especiales/hacia-donde-vamos-en-la-nueva-era-del-cliente

Alop, A. (2019). The Challenges of the Digital Technology Era for Maritime Education and Training. *2019 European Navigation Conference (ENC), Navigation Conference (ENC), 2019 European*. 10.1109/EURONAV.2019.8714176

AlSheibani, S., Cheung, Y., & Messom, C. (2018). Artificial intelligence adoption: AI-readiness at firm-level. *Proceedings of the 22nd Pacific Asia Conference on Information Systems - Opportunities and Challenges for the Digitized Society: Are We Ready? PACIS 2018*.

Al-Shura, M. S., Zabadi, A. M., Abughazaleh, M., & Alhadi, M. A. (2018). Critical success factors for adopting cloud computing in the pharmaceutical manufacturing companies. *Management and Economics Review*, *3*(2), 123–137. doi:10.24818/mer/2018.12-01

Alspaugh, S., Zokaei, N., Liu, A., Jin, C., & Hearst, M. A. (2018). Futzing and moseying: Interviews with professional data analysts on exploration practices. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 22–31. doi:10.1109/TVCG.2018.2865040 PMID:30136976

Alvarez-Meaza, I., Pikatza-Gorrotxategi, N., & Rio-Belver, R. M. (2020). Knowledge Sharing and Transfer in an Open Innovation Context: Mapping Scientific Evolution. *Journal of Open Innovation*, *68*(4), 1–23. doi:10.3390/joitmc6040186

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on big data in marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics*, *24*(1), 1–7. doi:10.1016/j.iedeen.2017.06.002

Amat Abreu, M., Ortega Tenezaca, D. B., & Yaguar Mariño, J. J. (2020). Determination of the degree of influence of the climatic factors of vulnerability of the agricultural sector with neutrosophic techniques [Determinación del grado de influencia de los factores climáticos de vulnerabilidad del sector agropecuario con técnicas neutrosóficas]. *Investigação Operacional*, *41*(5), 699–705.

Amatriain, X. (2013) *Beyond Data: from user information to business value through personalized recommendations and consumer science*. Paper presented at the CIKM'13, San Francisco, CA.

Ameen, N., Tarhini, A., Reppel, A., & Anand, A. (2020). Customer experiences in the age of artificial intelligence. *Computers in Human Behavior*, *114*, 106548. doi:10.1016/j.chb.2020.106548 PMID:32905175

Ameisen, E. (2018). *How to deliver on Machine Learning projects. A guide to the ML Engineering Loop.* Retrieved from https://blog.insightdatascience.com/how-to-deliver-on-machine-learning-projects-c8d82ce642b0

Amelia. (n.d.). *Amelia: The most human AI for the Enterprise*. https://amelia.com/amelia/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019). *Software engineering for machine learning: A case study.* Paper presented at the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). 10.1109/ICSE-SEIP.2019.00042

Amoah, A., Kwablah, E., Korle, K., & Offei, D. (2020). Renewable energy consumption in Africa: The role of economic well-being and economic freedom. *Sustainability Science*, *10*(32), 2–17.

Anderson, M. R. (2017). *After 75 Years, Isaac Asimov's Three Laws of Robotics Need Updating*. The Conversation.

Anderson, M., & Anderson, S. L. (Eds.). (2011). *Machine ethics*. Cambridge University Press. doi:10.1017/CBO9780511978036

Anderson, S. (2011). *Technologies of History*. Dartmouth College Press.

Anderson, S. L. (2007). *Asimov's "Three Laws of Robotics" and Machine Metaethics. Artificial Intelligence and Society, 22*.

Andrade, G., Ramalho, G., Gomes, A. S., & Corruble, V. (2006). Dynamic Game Balancing: An Evaluation of User Satisfaction. In *Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference* (p. 3–8). Menlo Park, CA: AAAI Press.

Andrade, J., & Fukuyama Sobata, M. (2020). Content aspects valued by users of travel blogs. Insights from brazilian travelers. *International Journal of Information Systems and Tourism*, *5*(1), 7–17.

Andras, P., Esterle, L., Guckert, M., Han, T. A., Lewis, P. R., Milanovic, K., Payne, T., Perret, C., Pitt, J., Powers, S. T., Urquhart, N., & Wells, S. (2018). Trusting intelligent machines: Deepening trust within socio-technical systems. *IEEE Technology and Society Magazine*, *37*(4), 76–83. doi:10.1109/MTS.2018.2876107

Angelis, J., & da Silva, E. R. (2019). Blockchain adoption: A value driver perspective. *Business Horizons*, *62*(3), 307–314. doi:10.1016/j.bushor.2018.12.001

Angelov, P., & Soares, E. (2020). Towards explainable deep neural networks (xDNN). *Neural Networks*, *130*, 185–194. doi:10.1016/j.neunet.2020.07.010 PMID:32682084

Ankit, K., Irsoy, O., Su, J., Bradbury, J., English, R., Pierce, B., Ondruska, P., Gulrajani, I., & Socher, R. (2016). *Ask Me Anything: Dynamic Memory Networks for Natural Language Processing*. ICML.

ANTAQ. (2019). *Port perfomance system: external user manual*. Agência Nacional de Transportes Aquaviários – ANTAQ.

Antonio, V. (2018, July 30). How AI is changing sales. *Harvard Business Review*. https://hbr.org/2018/07/how-ai-is-changing-sales

Anton, S. G., & Nucu, A. E. A. (2020). The effect of financial development on renewable energy consumption. A panel data approach. *Renewable Energy*, *147*, 330–338. doi:10.1016/j.renene.2019.09.005

Aoki, N. (2020). The importance of the assurance that "humans are still in the decision loop" for public trust in artificial intelligence: Evidence from an online experiment. *Computers in Human Behavior*, 106572.

Ardito, L., Petruzzelli, A. M., Panniello, U., & Garavelli, A. C. (2019). Towards industry 4.0: Mapping digital technologies for supply chain management-marketing integration. *Business Process Management Journal*, *25*(2), 323–346. doi:10.1108/BPMJ-04-2017-0088

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, *58*(2), 277–297. doi:10.2307/2297968

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51. doi:10.1016/0304-4076(94)01642-D

Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*, *82*(1), 150–169. doi:10.1006/obhd.2000.2893

Arnould, M., Morel, L., & Fournier, M. (2020). Developing a territorial diagnostic as part of a living lab process: Implementation to improve management and wood mobilization in small French private forest. *2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Engineering, Technology and Innovation (ICE/ITMC), 2020 IEEE International Conference On*. 10.1109/ICE/ITMC49519.2020.9198373

Arpteg, A., Brinne, B., Crnkovic-Friis, L., & Bosch, J. (2018). *Software engineering challenges of deep learning*. Paper presented at the 2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). 10.1109/SEAA.2018.00018

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Chatila, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. doi:10.1016/j.inffus.2019.12.012

Artaraz, M. (2002). Teoría de las tres dimensiones de desarrollo sostenible. *Ecosistemas (Madrid)*, *2002*(2). http//www.aeet.org/ecosistemas/022/informe1.htm

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 1014734. doi:10.1016/j.tele.2020.101473

Ashmore, R., Calinescu, R., & Paterson, C. (2019). *Assuring the machine learning lifecycle: Desiderata, methods, and challenges.* arXiv preprint arXiv:1905.04223.

Asimov, I. (1990). The Laws of Robotics. In Robot Visions. Academic Press.

Asimov, I. (n.d.). *Learn about Asimov's Three Laws of Robotics*. Britannica Publishing.

Asimov, I. (1950). *I, Robot*. Del Rey.

Asimov, I. (1957). *The Naked Sun*. Doubleday.

Asimov, I. (1991). *The Lord's Apprentice. In The Ultimate Frankenstein*. Manhanset House.

Asociación para el Desarrollo de la Experiencia de Cliente. (2017). La experiencia de cliente rentable. Libro colaborativo. Ed. DEC, Madrid 2017.

Asociación para el Desarrollo de la Experiencia de Cliente. (2019). *Behavioral Economics. VI Congreso Internacional sobre Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/eventos/congreso/congreso-2019/

Asociación para el Desarrollo de la Experiencia de Cliente. (2020). *The Experience Revolution. VII Congreso Internacional sobre Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/eventos/congreso/congreso-2020/

Assi, F., Isiksal, A. Z., & Tursoy, T. (2021). Renewable Energy Consumption, Financial Development, Environmental pollution, and Innovations in the ASEAN +3 group: Evidence from (P-ARDL) model. *Renewable Energy*, *165*(1), 689–700. doi:10.1016/j.renene.2020.11.052

Atlam, H. F., Azad, M. A., Alzahrani, A. G., & Wills, G. (2020). A Review of Blockchain in Internet of Things and AI. *Big Data and Cognitive Computing*, *4*(4), 28. doi:10.3390/bdcc4040028

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Attié, E., Gourdou, C., & Guédon, E. (2021). *Understanding consumer acceptance of neuromarketing* Paper presented at *International Marketing Trends Conference*, Venice.

Attwell, K., Lake, J., Sneddon, J., Gerrans, P., Blyth, C., & Lee, J. (2021). Converting the maybes: Crucial for a successful COVID-19 vaccination strategy. *PLoS One*, *16*(1), e0245907. doi:10.1371/journal.pone.0245907 PMID:33471821

Auh, S., Menguc, B., Katsikeas, C. S., & Jung, Y. S. (2019). When Does Customer Participation Matter? An Empirical Investigation of the Role of Customer Empowerment in the Customer Participation–Performance Link. *JMR, Journal of Marketing Research*, *56*(6), 1012–1033. doi:10.1177/0022243719866408

Autonomous ship project, key facts about YARA Birkeland. (2017). Retrieved from https://www.kongsberg.com/maritime/support/themes/autonomous-ship-project-key-facts-about-yara-birkeland/

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64. doi:10.103841586-018-0637-6 PMID:30356211

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *Management Information Systems Quarterly*, *30*(1), 13–28. doi:10.2307/25148715

Awang, Z., Afthanorhan, A., & Mamat, M. (2016). The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences*, *4*(1), 13.

Ayanso, A., Chatterjee, D., & Cho, D. I. (2011). E-government readiness index: A methodology and analysis. *Government Information Quarterly*, *28*(4), 522–532. doi:10.1016/j.giq.2011.02.004

Ayanso, A., Cho, D. I., & Lertwachara, K. (2010). The digital divide: Global and regional ict leaders and followers. *Information Technology for Development*, *16*(4), 304–319. doi:10.1080/02681102.2010.504698

Ayanso, A., Cho, D. I., & Lertwachara, K. (2014). Information and communications technology development and the digital divide: A global and regional assessment. *Information Technology for Development*, *20*(1), 60–77. doi:10.1080/02681102.2013.797378

Ayanso, A., & Lertwachara, K. (2015). An analytics approach to exploring the link between ict development and affordability. *Government Information Quarterly*, *32*(4), 389–398. doi:10.1016/j.giq.2015.09.009

Bachhofer, M., & Wildenberg, M. (2010). *FCMappers*. http://www.fcmappers.net

Backus, P., Janakiram, M., Mowzoon, S., Runger, C., & Bhargava, A. (2006). Factory Cycle-Time Prediction with a Data-Mining Approach. *Semiconductor Manufacturing. IEEE Transactions on.*, *19*(2), 252–258. doi:10.1109/TSM.2006.873400

Bacon, E., Williams, M. D., & Davies, G. H. (2019). Recipes for success: Conditions for knowledge transfer across open innovation ecosystems. *International Journal of Information Management*, *49*, 377–387. doi:10.1016/j.ijinfomgt.2019.07.012

Badri, M. A., & Alshare, K. (2008). A path analytic model and measurement of the business value of e-government: An international perspective. *International Journal of Information Management*, *28*(6), 524–535. doi:10.1016/j.ijinfomgt.2006.10.004

Baena Rojas, J. J., Cano Arenas, J. A., Jarrin Quintero, J. A., & Pérez Arroyave, H. R. (2014). Use of information and communication technologies for international negotiation: An advantage for Colombian companies? *Revista Ciencias Estratégicas, 22*(32), 279-294. Retrieved from https://www.redalyc.org/pdf/1513/151339264007.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bag, S., Pretorius, J. H. C., Gupta, S., & Dwivedi, Y. K. (forthcoming). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technological Forecasting and Social Change, 163*, 120420.

Bagherzadeh, M., Markovic, S., & Bogers, M. (2021). Managing Open Innovation: A Project-Level Perspective. *IEEE Transactions on Engineering Management, Engineering Management, IEEE Transactions on*. *IEEE Transactions on Engineering Management*, *68*(1), 301–316. doi:10.1109/TEM.2019.2949714

Bailis, P., Olukotun, K., Ré, C., & Zaharia, M. (2017). *Infrastructure for usable machine learning: The stanford dawn project.* arXiv preprint arXiv:1705.07538.

Baillargeon, N. (2018). L'éthique des voitures autonomes [The ethics of autonomous cars]. *Québec Science Magazine*. Accessed from: https://www.quebecscience.qc.ca/normand-baillargeon/ethique-des-voitures-autonomes/

Bain & Co, & Asociación para el Desarrollo de la Experiencia de Cliente. (2020). *V Informe de Madurez de Experiencia de Cliente.* Retrieved November 22, 2020, from https://asociaciondec.org/presentacion-de-informes/v-informe-de-madurez-de-experiencia-de-cliente/45773/

Bajracharya, K. (2018). *Data Science Glossary. Data Science.* Retrieved on 13 February 2021 from https://dimensionless.in/data-science-glossary/

Baker, V., Elliot, B., Sicular, S., Mullen, A., & Brethenoux, E. (2020). *Magic quadrant for cloud AI developer services*. Retrieved from https://www.gartner.com/en/documents/3981253/magic-quadrant-for-cloud-ai-developer-services

Bakkes, S. C. J., Spronck, P. & van Lankveld, G. (2012). Player behavioural modelling for video games. *Entertainment Computing, 3*, 71-79.

Balaguer. (2020). *¿Es posible analizar el ROI en la experiencia del cliente?* Retrieved March 15, 2021, from https://micliente.net/roi-en-la-experiencia-del-cliente/

Baldegger, R., Caon, M., & Sadiku, K. (2020). Correlation between Entrepreneurial Orientation and Implementation of AI in Human Resources Management. *Technology Innovation Management Review*, *10*(4), 72–79. doi:10.22215/timreview/1348

Ballesteros Riveros, D. P., & Ballesteros Silva, P. P. (2004). Competitive logistics and supply chain management. *Sciences et Techniques (Paris)*, 201–206. https://www.redalyc.org/pdf/849/84912053030.pdf

Balmer, J. M. T., & Greyser, S. A. (2002). Managing the multiple identities of the corporation. *California Management Review, 43*(3), 72-86. doi:10.2307/41166133

Balyan, R., McCarthy, K. S., & McNamara, D. S. (2020). Applying natural language processing and hierarchical machine learning approaches to text difficulty classification. *International Journal of Artificial Intelligence in Education*, *30*(3), 1–34. doi:10.100740593-020-00201-7

Ban, H. J., & Kim, H. S. (2019). Understanding customer experience and satisfaction through airline passengers' online review. *Sustainability*, *11*(15), 4066. doi:10.3390u11154066

Banini, G. A., & Bearman, R. A. (1998). Application of fuzzy cognitive maps to factors affecting slurry rheology.[. *International Journal of Mineral Processing*, *52*(4), 233–244. doi:10.1016/S0301-7516(97)00071-9

Bannister, F. (2007). The curse of the benchmark: An assessment of the validity and value of e-government comparisons. *International Review of Administrative Sciences*, *73*(2), 171–188. doi:10.1177/0020852307077959

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019a). Beyond accuracy: The role of mental models in human-ai team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (*Vol. 7*, No. 1, pp. 2-11). Academic Press.

Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., & Horvitz, E. (2019, July). Updates in human-ai teams: Understanding and addressing the performance/compatibility tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 2429–2437. doi:10.1609/aaai.v33i01.33012429

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. doi:10.1037/0022-3514.70.3.614 PMID:8851745

Bao, C., & Xu, M. (2019). Cause and effect of renewable energy consumption on urbanization and economic growth in China's provinces and regions. *Journal of Cleaner Production*, *231*, 483–493. doi:10.1016/j.jclepro.2019.05.191

Baran, T. (2020). Anket Formuyla Veri Toplamada Renk ve Cevap Alternatifi Sayısının Cevaplama Süresi ve İfade Cevaplama Sayısına Etkisi. *Akademisyen Yayınevi*. *Bölüm*, *5*, 97–116.

Barbic, J., D'Cruz, M., Latoschik, M. E., Slater, M., & Bourdot, P. (2017). Virtual Reality and Augmented Reality. In *14th EuroVR International Conference, EuroVR 2017 Laval*, *France*, December 12–14, 2017 *Proceedings.* Springer. 10.1007/978-3-319-72323-5

Bardakcı, A. (2009). Pazarlama Araştırmalarında Kullanılan Tutum Ölçeklerindeki Cevap Alternatifi Sayısına İlişkin Bir Literatür Taraması. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (4), 7–20.

Bargh, J. A. (1997). Automaticity of Everyday Life. *Advances in Social Cognition*, *10*, 1–61.

Barinaga, B. (2014). Digital distopy: Videogame. In Artecnología Augmented Knowledge and Accessibility. Complutense.

Barinaga, B. (2016). The video game and the game tradition.Education and shadows of the digital society. *Opción*, *32*, 9.

Barinaga, B. (2010). *Juego. Historia, teoría y práctica del diseño de videojuegos*. Alesia Games.

Barnette, J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*(3), 361–370. doi:10.1177/00131640021970592

Basak, S. K. (2017). A Framework on the Factors Affecting to Implement Maritime Education and Training System in Educational Institutions: A Review of the Literature. *Procedia Engineering*, *194*, 345–350. doi:10.1016/j.proeng.2017.08.155

Basdas, U., & Esen, M. F. (2021). *Review of Big Data Applications in Finance and Economics. In Handbook of Research on Engineering, Business, and Healthcare Applications of Data Science and Analytics*. IGI Global. doi:10.4018/978-1-7998-3053-5

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, *68*, 101–115. doi:10.1016/j.tourman.2018.02.020

Batterton, K., & Hale, K. (2017). The Likert Scale What It Is and How to Use It. *Phalanx*, *50*(2), 32–39.

Bauer, P., Stevens, B., & Hazeleger, W. (2021). A digital twin of Earth for the green transition. *Nature Climate Change*, *11*(2), 80–83. doi:10.103841558-021-00986-y

Bayle-Tourtoulou, A.-S., & Badoc, M. (2020). *The neuro-consumer: Adapting communication and marketing strategies for the consumer*. Routledge. doi:10.4324/9781003019978

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bazley, W. J., Cronqvist, H., & Mormann, M. (2021). Visual Finance: The Pervasive Effects of Red on Investor Behavior. *Management Science*. https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2020.3747

BBVA Research. (28 April 2016). *Peru: The retail sector.* Retrieved from https://www.bbvaresearch.com/publicaciones/peru-el-sector-retail/

BBVA. (2019). *Machine learning': ¿qué es y cómo funciona?* Retrieved November 25, 2020, from https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/

BBVA. (2020). *Blockchain y la inteligencia artificial: ¿La pareja ideal?* Retrieved December 23, 2020, from https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/blockchain-y-la-inteligencia-artificial-la-pareja-idealblockchain-y-la-inteligencia-artificial-la-pareja-ideal/

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293–1295. doi:10.1126cience.275.5304.1293 PMID:9036851

Beck, M., & Libert, B. (2019). The machine learning race is really a data race. *MIT Sloan Management Review*. Available at https://sloanreview.mit.edu/article/the-machine-learningrace-is-really-a-data-race/

Becker, M. H. (1974). The health belief model and sick role behavior. *Health Education Monographs*, *2*(4), 409–419. doi:10.1177/109019817400200407

Bednar, K., Spiekermann, S., & Langheinrich, M. (2019). Engineering Privacy by Design: Are engineers ready to live up to the challenge? *The Information Society*, *35*(3), 122–142. doi:10.1080/01972243.2019.1583296

Belanche, D., Casaló, L. V., & Flavián, C. (2019). Artificial Intelligence in FinTech: Understanding robo-advisors adoption among customers. *Industrial Management & Data Systems*, *119*(7), 1411–1430. Advance online publication. doi:10.1108/IMDS-08-2018-0368

Belani, H., Vukovic, M., & Car, Ž. (2019). *Requirements engineering challenges in building AI-based complex systems.* Paper presented at the 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW). 10.1109/REW.2019.00051

Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., & Mojsilović, A. (2019). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development, 63*(4/5), 1-4.

Bell, P. (2004). Monster or Machine? *Nature*.

Benson, C., D. S., P. & Colling, A. (2018, June). *A Quantitative Analysis of Possible Futures of Autonomous Transport*. Academic Press.

Benton, W. C. (2020). Machine learning systems and intelligent applications. *IEEE Software*, *37*(04), 43–49. doi:10.1109/MS.2020.2985224

Benzidia, S., Makaoui, N., & Bentahar, O. (2020). The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. *Technological Forecasting and Social Change*, *165*, 120557. doi:10.1016/j.techfore.2020.120557

Berezina, K., Ciftci, O., & Cobanoglu, C. (2019). Robots, artificial intelligence, and service automation in restaurants. In *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*. Emerald Publishing Limited. doi:10.1108/978-1-78756-687-320191010

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Berger, J., & Milkman, K. (2012). What makes online content viral? *JMR, Journal of Marketing Research*, *49*(2), 192–205. doi:10.1509/jmr.10.0353

Berman, S. J., Kesterson-Townes, L., Marshall, A., & Srivathsa, R. (2012). How cloud computing enables process and business model innovation. *Strategy and Leadership*, *40*(4), 27–35. doi:10.1108/10878571211242920

Beşikçi, E. B., Arslan, O., Turan, O., & Ölçer, A. I. (2016, February). An artificial neural network-based decision support system for energy-efficient ship operations. *Computers & Operations Research*, *66*, 393–401. doi:10.1016/j.cor.2015.04.004

Beuscart, J. S., Mellet, K., & Trespeuch, M. (2016). Reactivity without legitimacy? Online consumer reviews in the restaurant industry. *Journal of Cultural Economics*, *9*(5), 458–475. doi:10.1080/17530350.2016.1210534

Bhatia, R. (2016, December 12). Say hello to Simi – a chatbot that makes your cocktail. *Analytics India Magazine*. https://analyticsindiamag.com/say-hello-simi-chatbot-makes-cocktail/

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly: Management*. *Information Systems*, *25*(3), 351. Advance online publication. doi:10.2307/3250921

Bigman, Y. E., & Gray, K. (2020). Life and Death Decisions of Autonomous Vehicles. *Nature*, *579*(7797), E1–E2. doi:10.103841586-020-1987-4 PMID:32132695

Bilgen, B., & Ozkarahan, I. (2007, June). A mixed-integer linear programming model for bulk grain blending and shipping. *International Journal of Production Economics*, *107*(2), 555–571. doi:10.1016/j.ijpe.2006.11.008

Bilgihan, A., Kandampully, J., & Zhang, T. (2015). Towards a unified customer experience in online shopping environments. Antecedents and outcomes. *International Journal of Quality and Service Sciences*, *8*(1), 102-119. doi:10.1108/IJQSS-07-2015-0054

Bilgihan, A., Barreda, A., Okumus, F., & Nusair, K. (2016). Consumer perception of knowledge sharing in travel related online social networks. *Tourism Management*, *52*, 287–296. doi:10.1016/j.tourman.2015.07.002

Bing Liu, Hakkani-Tur, & Heck. (n.d.). *Dialogue Learning with Human Teaching and Feedback in End-to-End Trainable Task-Oriented Dialogue Systems*. Academic Press.

Bird, S., Dudík, M., Edgar, R., Horn, B., Lutz, R., Milan, V., Sameki, M., Wallach, H., & Walker, K. (2020). *Fairlearn: A toolkit for assessing and improving fairness in AI*. Retrieved from https://www.microsoft.com/en-us/research/uploads/prod/2020/05/Fairlearn_WhitePaper-2020-09-22.pdf

Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. *International Journal of Exercise Science*, *8*(3), 297. PMID:27182418

Blank, D. (2006). *Robotics and Asimov's Three Laws. In The Frankenstein Complex and Asimov's Three Laws*. Association of the Advancement of Artificial Intelligence.

Blum, P. R., (2016). Robots, Slaves, and the Paradox of the Human Condition. In *Isaac Asimov's Robot Stories*. Czech Science Foundation at the project Between Renaissance and Baroque Philosophy and Knowledge of the Czech Lands within the Wider European Context.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143. doi:10.1016/S0304-4076(98)00009-8

Blut, M., Beatty, S. E., Evanschitzky, H., & Brock, C. (2014). The impact of service characteristics on the switching costs–customer loyalty link. *Journal of Retailing*, *90*(2), 275–290. doi:10.1016/j.jretai.2014.04.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bogachov, S., Kwilinski, A., Miethlich, B., Bartosova, V., & Gurnak, A. (2020). Artificial intelligence components and fuzzy regulators in entrepreneurship development. *Entrepreneurship and Sustainability Issues*, *8*(2), 487–499. doi:10.9770/jesi.2020.8.2(29)

Bogers, M., Chesbrough, H., Heaton, S., & Teece, D. J. (2019). Strategic Management of Open Innovation: A Dynamic Capabilities Perspective. *California Management Review*, *62*(1), 77–94. doi:10.1177/0008125619885150

Bohannon, J., Koch, D., Homm, P., & Driehaus, A. (2015, December). Chocolate with high cocoa content as a weight-loss accelerator. *International Archives of Medicine*, *8*(55), 1–8.

Boletín Oficial del Estado. (2018). *Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*. Retrieved November 18, 2020, from https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673&p=20181206&tn=2

Bolstrom, N., & Yudkowsky, E. (2011). *The Ethics of Artificial Intelligence*. Cambridge University Press.

Bond, S. R. (2002). *Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice.* Center for Microdata Methods and Practice Working Papers.

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The Social Dilemma of Autonomous Vehicles. *Science*, *352*(6293), 1573–1576. doi:10.1126cience.aaf2654 PMID:27339987

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, *50*(2), 1–5.

Boone, J. M. (1997). Hotel-restaurant co-branding: A preliminary study. *The Cornell Hotel and Restaurant Administration Quarterly*, *38*(5), 34–43. doi:10.1016/S0010-8804(97)86581-0

Bornstein, R. F., & Pittman, T. S. (1992). *Perception Without Awareness*. The Guilford Press.

Bosch, J., Crnkovic, I., & Olsson, H. H. (2020). *Engineering AI systems: A research agenda.* arXiv preprint arXiv:2001.07522.

Boston Consulting Group & DEC. (2018). Informe BCG-DEC. In *Reinventarse: El gran reto de la Experiencia de Cliente. Estudio integral sobre la Experiencia de Cliente en España Evolución 2016-2018*. Retrieved November 28, 2020, from https://asociaciondec.org/informes-dec/informe-dec_bcg-reinventarse-el-gran-reto-de-la-experiencia-de-cliente/35680/

Bouguessa, M. (2011). A Practical Approach for Clustering Transaction Data. In P. Perner (Ed.), *Machine Learning and Data Mining in Pattern Recognition* (Vol. 6871, pp. 265–279). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-23199-5_20

Bourque, P., & Fairley, R. (2014). Guide to the software engineering body of knowledge (SWEBOK (R)): Version 3.0. IEEE Computer Society Press.

Bozonier, J. (2015). *Test-driven machine learning*. Packt Publishing Ltd.

Brantinger, P. (1980). The Gothic Origins of Science Fiction. *NOVEL: A Forum on Fiction, 14*(1), 30–43.

Braschler, M., Stadelmann, T., & Stockinger, K. (2019). *Applied data science : lessons learned for the data-driven business. Applied data science lessons learned for the data-driven business*. Springer., doi:10.1007/978-3-030-11821-1

Brazil Country Review. (2021). Retrieved from https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=149896693&site=eds-live

Breck, E., Cai, S., Nielsen, E., Salib, M., & Sculley, D. (2016). *What's your ML Test Score? A rubric for ML production systems*. Paper presented at the Reliable Machine Learning in the Wild - NIPS 2016 Workshop.

Brennan, R., Canning, L., & McDowell, R. (2014). *Business-to-Business Marketing* (3rd ed.). Sage.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. doi:10.1016/j.jss.2006.07.009

Breuel, C. (2020, Jan 3). *MLOps: Machine learning as an engineering discipline*. Retrieved from https://towardsdatascience.com/ml-ops-machine-learning-as-an-engineering-discipline-b86ca4874a3f

Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D., & Weinstein, N. D. (2007). Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychology*, *26*(2), 136–145. doi:10.1037/0278-6133.26.2.136 PMID:17385964

Brill, T. M., Munoz, L., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*(15-16), 1401–1436. doi:10.1080/0267257X.2019.1687571

Brock, D. C. (2018). Learning from Artificial Intelligence's Previous Awakenings: The History of Expert Systems. *AI Magazine*, *39*(3), 3–15. doi:10.1609/aimag.v39i3.2809

Brooks, R. A. (1991). *Intelligence without Reason*. Massachusetts Institute of Technology Artificial Intelligence Laboratory.

Brooks, C., Hoepner, A. G. F., McMillan, D., Vivian, A., & Simen, C. W. (2019). Financial data science: The birth of a new financial research paradigm complementing econometrics? *European Journal of Finance*, *25*(17), 1627–1636. doi:10.1080/1351847X.2019.1662822

Brooks, M. R., Button, K., & Nijkamp, P. (2002). *Maritime Transport*. Edward Elgar Pub.

Brown, J. D. (2011). Likert items and scales of measurement. *Statistics*, *15*(1), 10–14.

Brundtland, G. H. (1987). *Report of the World Commission on Environment and Development: Our Common Future ('The Brundtland Report').* doi:10.9774/GLEAF.978-1-907643-44-6_12

BrynjolfssonE.HittL.KimH. (2011). Strength in numbers: how does data-driven decision making affect firm performance? https://ssrn.com/abstract=1819486 doi:10.2139srn.1819486

Bryson, S. (1992). Measurement and calibration of static distortion of position data from 3D trackers. *SPIE Proceedings Vol. 1669: Stereoscopic Displays and Applications III*.

Buck, B., Espinach, L., & Söderberg, S. (2014). *GRI G4 Guidelines and ISO 26000:2010 How to use the GRI G4 Guidelines and ISO 26000 in conjunction*. Retrieved from https://www.iso.org/files/live/sites/isoorg/files/archive/pdf/en/iso-gri-26000_2014-01-28.pdf

Buck, J., & Lee, E. (1993). Scheduling dynamic dataflow graphs with bounded memory using the token flow model. *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1, 429-432. 10.1109/ICASSP.1993.319147

Bughin, J., Cincera, M., Reykowska, D., Żyszkiewicz, M., & Ohme, R. (2020). *Perceptive risk clusters of European citizens and NPI compliance in face of the covid-19 pandemics.* Working Papers ECARES 2020-51, ULB Université libre de Bruxelles.

Bughin, J., Hazan, E., Ramaswamy, S., Allas, T., Dahlström, P., Henke, N., & Trench, M. (2017). Artificial intelligence: the next digital frontier? *McKinsey Global Institute*. https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx

Bughin, J. (2016). Big data, Big bang? *Journal of Big Data*, *3*(1), 1–14. doi:10.118640537-015-0014-3

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

BughinJ.CinceraM. (2021). *How Institutional Actions before Vaccine Affects Vaccination Intention Later: Prediction Via Machine Learning*. https://ssrn.com/abstract=3824550 doi:10.2139srn.3824550

BughinJ.CinceraM.ReykowskaD.ŻyszkiewiczM.OhmeR. (2021). COVID-19 Endemism and the Control Skeptics. https://ssrn.com/abstract=3785230 doi:10.2139srn.3785230

Buhalis, D., & Leung, R. (2018). Smart hospitality—Interconnectivity and interoperability towards an ecosystem. *International Journal of Hospitality Management*, *71*, 41–50. doi:10.1016/j.ijhm.2017.11.011

Burke, R., & Leykin, A. (2014). *Identifying the Drivers of Shopper Attention*. Engagement, and Purchase., doi:10.1108/S1548-643520140000011006

Burkov, A. (2020). *Machine learning engineering*. True Positive Inc.

Burns, A., & Bush, R. (2000). *Marketing Research* (3rd ed.). Prentice Hall.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 2053951715622512. doi:10.1177/2053951715622512

Buttle, F., & Maklan, S. (2019). Customer relationship management: concepts and technologies (4th ed.). Routledge, Taylor & Francis Group. doi:10.4324/9781351016551

Buxbaum, P. (2020). Pandemic developments will drive future supply-chain improvements. *American Journal of Transportation*.

Bybee, R. W. (1991). Planet Earth in Crisis: How Should Science Educators Respond? *The American Biology Teacher*, *53*(3), 146–153. doi:10.2307/4449248

Byrne, J., Hughes, K., Rickerson, W., & Kurdgelashvili, L. (2007). American policy conflict in the greenhouse: Divergent trends in federal, regional, state, and local green energy and climate change policy, *Energy*. *Pol*, *35*(9), 4555–4573. doi:10.1016/j.enpol.2007.02.028

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. doi:10.1037/0033-2909.125.3.367

Caceres Cabana, Y. D., Gonzalez, S., & Rivera Medina, J. (2020). Alpaca meat: constraints and reinforcements in consumption. In Anthropological approaches to understanding consumption patterns and consumer behaviour. IGI Global. doi:10.4018/978-1-7998-3115-0.ch016

Cacho, J. L., Marques, L., & Nascimento, Á. (2020). Customer-Oriented Global Supply Chains: Port Logistics in the Era of Globalization and Digitization. In V. Chkoniya, A. O. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 82–103). IGI Global. doi:10.4018/978-1-7998-3115-0.ch005

CAF Development Bank of Latin America. (2020). *The state of Latin America's digitalization in the face of the COVID-19 pandemic* (Vol. 19). Retrieved from Observatorio CAF del Ecosistema Digital: https://scioteca.caf.com/handle/123456789/1540

CAF Development Bank of Latin America. (2020). The state of Latin America's digitalization in the face of the COVID-19 pandemic. *CAF Observatory of the Digital Ecosystem, 19*(3), 17. Retrieved from https://scioteca.caf.com/handle/123456789/1540

Cai, S., Bileschi, S., Nielsen, E. D., & Chollet, F. (2020). *Deep learning with JavaScript*. Manning.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Calders, T., & Žliobaitė, I. (2013). Why unbiased computational processes can lead to discriminative decision procedures. In *Discrimination and privacy in the information society* (pp. 43–57). Springer. doi:10.1007/978-3-642-30487-3_3

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. doi:10.1109/T-AFFC.2010.1

Camacho, D., Panizo-Lledot, A., Bello-Orgaz, G., & Gonzalez-Pardo, C. E. (2020). The four dimensions of social network analysis: an overview of research methods, applications, and software tools. Information Fusion, 63, 88–120. doi:10.1016/j.inffus.2020.05.009

Camarinha-Matos, L. M. (2012). *Scientific research methodologies and techniques- Unit 2: Scientific method. Unit 2: Scientific methodology. PhD program in electrical and computer engineering*. Uninova.

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons*, *63*(2), 183–193. doi:10.1016/j.bushor.2019.11.003

Cao, L. (2017). Data science: A comprehensive overview. *ACM Comput. Surv. 50*(3). doi:10.1145/3076253

Cao, K., Liu, E., & Jain, A. K. (2014). Segmentation and Enhancement of Latent Fingerprints: A Coarse to Fine RidgeStructure Dictionary. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(9), 1847–1859. doi:10.1109/TPAMI.2014.2302450 PMID:26352236

Capgemini. (2005). *Online availability of public services: How is europe progressing?* Retrieved from http://www.umic.pt/images/stories/publicacoes/2004[1].01%20-%20Web-based%20Survey%20on%20Electronic%20Public%20Services%20-%20Fourth%20measurement.pdf

Capgemini. (2006). *Online availability of public services: How is europe progressing?* Retrieved from https://www.ut.is/media/Skyrslur/CapGemini_2006.pdf

Capgemini. (2020). *Informe mundial del seguro 2020 de Capgemini*. Retrieved December 15, 2020, from https://www.capgemini.com/es-es/news/informe-mundial-del-seguro-2020-de-capgemini/

Caputo, A., Marzi, G., & Pellegrini, M. (2016). The Internet of Things in manufacturing innovation processes: Development and application of a conceptual framework. *Business Process Management Journal*, *22*(2), 383–402. doi:10.1108/BPMJ-05-2015-0072

Carbone, V., & De Martino, M. (2003). The changing role of ports in supply-chain management: An empirical analysis. *Maritime Policy & Management*, *30*(4), 305–320. doi:10.1080/0308883032000145618

Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, *42*(12), 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x PMID:19120943

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106–116. doi:10.3844/jssp.2007.106.116

Carley, K., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models.[. *Social Forces*, *70*(3), 601–636. doi:10.2307/2579746

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, *51*(1), 341–377. doi:10.100711042-010-0660-6

Carr, S. (2021). *How Many Ads Do We See A Day In 2021?* https://ppcprotect.com/how-many-ads-do-we-see-a-day/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Frontiers in Behavioral Neuroscience*, *3*, 21. doi:10.3389/neuro.08.021.2009 PMID:19753142

Carù, A., & Cova, B. (2013). *Consuming experiences: An introduction. In Consuming experience*. Routledge. doi:10.4324/9780203390498

Casaca, A. C. P., & Loja, M. A. R. (2020). 2019 World of Shipping Portugal. An International Research Conference on Maritime Affairs editorial "Leading the shipping industry into the future." *Journal of Shipping & Trade*, *5*(1), 1–4. doi:10.118641072-020-00082-y

Casco, A. R. (2020). Effects of the COVID-19 pandemic on consumer behavior. *INNOVARE Journal of Science and Technology, 9*(2).

Cassar, M. L., Caruana, A., & Konietzny, J. (2020). Wine and satisfaction with fine dining restaurants: An analysis of tourist experiences from user generated content on TripAdvisor. *Journal of Wine Research*, 1–16.

Castells, M. (1996). La era de la información. Economía, sociedad y cultura. Siglo XXI.

Castelo, N. (2019). *Blurring the line between human and machine: Marketing artificial intelligence* (Doctoral dissertation). Columbia University Academic Commons. doi:10.7916/d8-k7vk-0s40

Castelo, N., Bos, M., & Lehman, D. (2018). *Consumer adoption of algorithms that blur the line between human and machine*. Graduate School of Business: Columbia University Working Paper.

Castelo, N., & Ward, A. (2016). Political affiliation moderates attitudes towards artificial intelligence. In P. Moreau & S. Puntoni (Eds.), *NA - Advances in Consumer Research* (pp. 723–723). Association for Consumer Research.

Castrillón-Muñoz, A. J., Infante-Moro, A., Zuñiga-Collazos, A., & Martínez-López, F. J. (2019). University Spin-Off: A Literature Review for Their Application in Colombia. *Journal of Environmental Management and Tourism*, *10*(1), 73–86. doi:10.14505//jemt.v10.7(33).08

Castrillón-Muñoz, A. J., Infante-Moro, A., Zuñiga-Collazos, A., & Martínez-López, F. J. (2020a). Capacities of the research groups at unicauca (Colombia) to develop spin-off-type undertakings. *Journal of Technology Management & Innovation*, *15*(1), 64–75. doi:10.4067/S0718-27242020000100064

Castrillón-Muñoz, A. J., Infante-Moro, A., Zúñiga-Collazos, A., & Martínez-López, F. J. (2020b). Generation of technological-based spin offs based on I+D+i results from the research groups of the University of Cauca Colombia [Generación de empresas derivadas de base tecnológica (Spin Offs), a partir de los resultados de I+ D+ i de los grupos de investigación de la Universidad del Cauca, Colombia]. *Información Tecnológica*, *31*(1), 67–78.

Cavoukian, A. (2012). Privacy by design: origins, meaning, and prospects for assuring privacy and trust in the information era. In G. O. M. Yee (Ed.), *Privacy protection measures and technologies in business organizations: aspects and standards* (pp. 170–208). IGI Global. doi:10.4018/978-1-61350-501-4.ch007

Centomo, S., Panado, M., & Fummi, F. (2018). Cyber-Physical Systems Integration in a Production Line Simulator. *2018 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)*, 237-242. 10.1109/VLSI-SoC.2018.8644836

Central Reserve Bank of Peru. (2016). *Informe Económico y Social Región Arequipa.* Retrieved from https://www.bcrp.gob.pe/docs/Proyeccion-Institucional/Encuentros-Regionales/2016/arequipa/ies-arequipa-2016.pdf

CepymeNews. (2020). *14 estadísticas sobre la gestión de la experiencia del cliente.* Retrieved November 16, 2020, from https://cepymenews.es/gestion-experiencia-cliente/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cerdán, Á. L. M. (2006). Análisis de la adopción de tecnologías colaborativas en Pymes. *Revista de economía y empresa*, *24*(56), 51-66.

Chan, A. P. H., & Tung, V. W. S. (2019). Examining the effects of robotic service on brand experience: The moderating role of hotel segment. *Journal of Travel & Tourism Marketing*, *36*(4), 458–468. Advance online publication. doi:10.1080/10548408.2019.1568953

Chang, K. (2020). Artificial intelligence in personnel management: The development of APM model. *The Bottom Line (New York, N.Y.)*, *33*(4), 377–388. doi:10.1108/BL-08-2020-0055

Chang, K. C., Kuo, N. T., Hsu, C. L., & Cheng, Y. S. (2014). The impact of website quality and perceived trust on customer purchase intention in the hotel sector: Website brand and perceived value as moderators. *International Journal of Innovation, Management and Technology*, *5*(4), 255. doi:10.7763/IJIMT.2014.V5.523

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. Academic Press.

Chapman, V., & Hunicke, R. (2004). *AI for Dynamic Difficulty Adjustment in Games*. Academic Press.

Charmetant, E. (2018). La neuroéthique: Histoire, actualité et prospective [Neuroethics: history, present and future]. *Transversalites*, *146*(3), 45–58. doi:10.3917/trans.146.0045

Chaturvedi, A., Carroll, J. D., Green, P. E., & Rotondo, J. A. (1997). A feature-based approach to market segmentation via overlapping k-centroids clustering. *JMR, Journal of Marketing Research*, *34*(3), 370–377. doi:10.1177/002224379703400306

Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing*, *65*(2), 81–93. doi:10.1509/jmkg.65.2.81.18255

Chen, C.-H., Khoo, L. P., Chong, Y. T., & Yin, X. F. (2014, May). Knowledge discovery using genetic algorithm for maritime situational awareness. *Expert Systems with Applications*, *41*(6), 2742–2753. doi:10.1016/j.eswa.2013.09.042

Chen, C., Härdle, W. K., & Unwin, A. (2007). *Handbook of data visualization*. Springer Science & Business Media.

Chen, C., & Phou, S. (2013). A closer look at destination: Image, personality, relationship and loyalty. *Tourism Management*, *36*, 269–278. doi:10.1016/j.tourman.2012.11.015

Chen, F., Feng, J., Jain, A. K., Zhou, J., & Zhang, J. (2011). Separating overlapped fingerprints. *IEEE Transactions on Information Forensics and Security*, *6*(2), 346–359. doi:10.1109/TIFS.2011.2114345

Chen, H., & Boning, D. (2017). Online and Incremental Machine Learning Approaches for IC Yield Improvement. *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. 10.1109/ICCAD.2017.8203857

Chiambaretto, P., & Gurău, C. (2017). David by Goliath: What is co-branding and what is in it for SMEs. *International Journal of Entrepreneurship and Small Business*, *31*(1), 103–122. doi:10.1504/IJESB.2017.083805

Chkoniya, V. (2021). *How Harvard's Case Method Can Help to Overcome the Challenges of Applied Data Science Education*. Academic Press.

Chkoniya, V., & Mateus, A. (2019). Digital Category Management: How Technology Can Enable the Supplier-Retailer Relationship. In Smart Marketing With the Internet of Things (pp. 139-163). IGI Global.

Chkoniya, V., Madsen, A. O., & Bukhrashvili, P. (2020). *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior*. IGI Global. doi:10.4018/978-1-7998-3115-0

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chopra, K. (2019). Indian shopper motivation to use artificial intelligence: Generating Vroom's expectancy theory of motivation using grounded theory approach. *International Journal of Retail & Distribution Management*, *47*(3), 331–347. Advance online publication. doi:10.1108/IJRDM-11-2018-0251

Cho, S. W. (2020). Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. *The Econometrics Journal*, *23*(3), 323–334. doi:10.1093/ectj/utaa025

Chou, K. L., Amick, M. M., Brandt, J., Camicioli, R., Frei, K., Gitelman, D., Goldman, J., Growdon, J., Hurtig, H. I., Levin, B., Litvan, I., Marsh, L., Simuni, T., Tröster, A. I., & Uc, E. Y. (2010). A Recommended Scale for Cognitive Screening in Clinical Trials of Parkinson's Disease. *Movement Disorders*, *25*(15), 2501–2507. doi:10.1002/mds.23362 PMID:20878991

Christian, L. M., Parsons, N. L., & Dillman, D. A. (2009). Designing scalar questions for Web surveys. *Sociological Methods & Research*, *37*(3), 393–425. doi:10.1177/0049124108330004

Churchill, G. A., Jr. (1999). Marketing Research Methodological Foundations (7th ed.). The Dryden Press International Edition.

Chyung, S. Y., Barkin, J. R., & Shamsy, J. A. (2018). Evidence-Based Survey Design: The Use of Negatively Worded Items in Surveys. *Performance Improvement*, *57*(3), 16–25. doi:10.1002/pfi.21749

Chyung, S. Y., Hutchinson, D., & Shamsy, J. A. (2020). Evidence-Based Survey Design: Ceiling Effects Associated with Response Scales. *Performance Improvement*, *59*(6), 6–13. doi:10.1002/pfi.21920

Chyung, S. Y., Kennedy, M., & Campbell, I. (2018). Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert-Type Response Options. *Performance Improvement*, *57*(9), 9–16. doi:10.1002/pfi.21800

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement*, *56*(10), 15–23. doi:10.1002/pfi.21727

Chyung, S. Y., Swanson, I., Roberts, K., & Hankinson, A. (2018). Evidence-based survey design: The use of continuous rating scales in surveys. *Performance Improvement*, *57*(5), 38–48. doi:10.1002/pfi.21763

Ciborra, C. (2005). Interpreting e-government and development: Efficiency, transparency or governance at a distance? *Information Technology & People*, *18*(3), 260–279. doi:10.1108/09593840510615879

Ciccarelli, S. K., & Meyer, G. E. (2006). *Psychology*. Prentice Hall.

Cisco. (2020). *Cisco Annual Internet Report (2018–2023).* Retrieved November 21, 2020, from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

Claeys, A., & Vialatte, J.-S. (2012). Office parlementaire d'évaluation des choix scientifiques et technologiques. L'impact et les enjeux des nouvelles technologies d'exploration et de thérapie du cerveau [Parliamentary Office for the Evaluation of Scientific and Technological Choices. The impact and challenges of new technologies for brain exploration and therapy]. *Rapport par Claeys et Vialatte, députés. Paris. Assemblée Nationales et Sénat*, *2012*, 25.

Clancy, L. (2019). *Functional programming reaches for stardom in finance*. CME Group. Accessed and reviewed in April 2021, https://www.risk.net/risk-management/6395366/functional-programming-reaches-for-stardom-in-finance

Clark, K. R., Moses, E., & Leslie, K. R. (2021). System N: A Multimodal Approach to Understanding Consumer Decision Making. *Journal of Consumer Psychology*.

Cliffydcw (Producer). (2020). *Own work.* Retrieved from https://commons.wikimedia.org/w/index.php?curid=19054763

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cluetrain. (1999). *95 Theses*. Retrieved November 16, 2020, from https://www.cluetrain.com/; http://tremendo.com/cluetrain/

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, *40*(1), 79–96. doi:10.1093/aepp/ppx056

Codara, L. (1998). *Le Mappe Cognitive*. Carocci Editore.

Cohen, H. (2013). *Los 4 momentos de la verdad en el marketing*. Retrieved December 13, 2020, from https://www.marketingdirecto.com/actualidad/checklists/los-4-momentos-de-la-verdad-en-el-marketing

Cohen, J. B., Pham, M.T., & Andrade, E.B. (2006): The nature and role of affect in consumer behavior. In Handbook of Consumer Psychology. Lawrence Erlbaum.

Columbus, L. (2019). *10 Ways Machine Learning Is Revolutionizing Manufacturing in 2019*. https://www.forbes.com/sites/louiscolumbus/2019/08/11/10-ways-machine-learning-is-revolutionizing-manufacturing-in-2019/?sh=7e5167012b40

Consoli, S., Reforgiato Recupero, D., & Saisana, M. (Eds.). (2021). Data Science for Economics and Finance. Methodologies and Applications. Springer International Publishing. doi:10.1007/978-3-030-66891-4

Constanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds we Need*. MIT Press. doi:10.7551/mitpress/12255.001.0001

Consumer 360 and DATUM International. (2020). *Customer Experience and COVID-19 Comercio, Bancos y Farmacias protocols.* Retrieved from https://www.datum.com.pe/new_web_files/files/pdf/2020%20Experiencia%20del%20Cliente%20y%20Protocolos%20Covid-19.pdf

Contissa, G., Lagioia, F., & Sartor, G. (2017). The ethical knob: Ethically-customizable automated vehicles and the law. *Artificial Intelligence and Law*, *25*(3), 365–378. doi:10.100710506-017-9211-z

Copeland, B. J., & Produfoot, D. (2004). *The Computer, Artificial Intelligence, and the Turing Test. In Alan Turing: Life and Legacy of a Great Thinker*. Springer.

Corcos, A., & Pannequin, F. (2013). Neuroeconomics, decision-making and rationality. *Économie et institutions, 16*. . doi:10.4000/ei.74

Corrocher, N., & Cappa, E. (2020). The Role of public interventions in inducing private climate finance: An empirical analysis of the solar energy sector. *Energy Policy*, *147*, 111787. doi:10.1016/j.enpol.2020.111787

Cosic, R., Shanks, G., & Maynard, S. B. (2015). A business analytics capability framework. *AJIS. Australasian Journal of Information Systems*, *19*, S5–S19. doi:10.3127/ajis.v19i0.1150

Courbet, D., & Benoit, D. (2013). Neurosciences au service de la communication commerciale: manipulation et éthique - Une critique du neuromarketing [Neuroscience at the service of commercial communication: manipulation and ethics - A Critique of Neuromarketing]. *Études de communication - Langages, information, médiations*. *Université de Lille*, *40*, 28–42.

Coussement, K., Lessmann, S., & Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, *95*, 27–36. doi:10.1016/j.dss.2016.11.007

Crawford, C. (2005). *On Interactive Storytelling*. New Riders.

Crews, D. E., & Zavotka, S. (2006). Aging, Disability, and Frailty: Implications for Universal Design. *Physiol Anthropol.*, *25*(1), 113–118. doi:10.2114/jpa2.25.113 PMID:16617216

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cross, J., Meadow, G., & Com, I., & Jostaphot. (2017, September). Autonomous ships 101. *Journal of Ocean Technology*, *12*, 23–27.

Crupi, A., Del Sarto, N., Di Minin, A., Gregori, G. L., Lepore, D., Marinelli, L., & Spigarelli, F. (2020). The digital transformation of SMEs–a new knowledge broker called the digital innovation hub. *Journal of Knowledge Management*, *24*(6), 1263–1288. doi:10.1108/JKM-11-2019-0623

Cruz-Neira, C. (1993). Virtual reality overview. *SIGGRAPH 93 International Conference on Computer Graphics and Interactive Techniques*.

Cubilla-Montilla, M., Rodriguez, C., & Ortega, E. (2015). Análisis de los indicadores de sostenibilidad del Global Reporting Initiative. Una mirada desde el Biplot Logístico. *Revista Centros*, *3*, 96–114.

Cubric, M. (2020). Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study. *Technology in Society*, *62*, 101257. Advance online publication. doi:10.1016/j.techsoc.2020.101257

Cullen, R. (2001). Tackling the digital divide. *Online Information Review*, *25*(5), 311–320. doi:10.1108/14684520110410517

Curia, L., & Lavalle, A. (2011). Decision strategies in dynamic systems - applying fuzzy cognitive maps application to a socio-economic example [Estrategias de decisión en sistemas dinámicos - aplicando mapas cognitivos difusos aplicación a un ejemplo socio-económico]. *Journal of Information Systems and Technology Management*, *8*, 663–680. doi:10.4301/S1807-17752011000300008

Custers, B. (2013). Data dilemmas in the information society: Introduction and overview. In *Discrimination and Privacy in the Information Society* (pp. 3–26). Springer. doi:10.1007/978-3-642-30487-3_1

Czermański, E., Cirella, G. T., Oniszczuk-Jastrząbek, A., Pawłowska, B., & Notteboom, T. (2021). An Energy Consumption Approach to Estimate Air Emission Reductions in Container Shipping. *Energies*, *14*(278), 278. doi:10.3390/en14020278

Daellenbach, H., McNickle, D., & Dye, Sh. (2012). *Management Science. Decision-making through systems thinking* (2nd ed.). Plagrave Macmillian.

Dafoe, A. (2018). AI governance: a research agenda. Governance of AI Program, Future of Humanity Institute, University of Oxford.

Dahlberg, M., & Johansson, E. (2000). An examination of the dynamic behaviour of local governments using GMM bootstrapping methods. *Journal of Applied Econometrics*, *15*(4), 401–416. doi:10.1002/1099-1255(200007/08)15:4<401::AID-JAE564>3.0.CO;2-G

Dahlstrom, R., & Dato-on, M. C. (2004). Business-to-business antecedents to retail co-branding. *Journal of Business-To-Business Marketing*, *11*(3), 1–22. doi:10.1300/J033v11n03_01

Dai, L., Hu, H., & Zhang, D. (2015, October). An empirical analysis of freight rate and vessel price volatility transmission in global dry bulk shipping market [English Edition]. *Journal of Traffic and Transportation Engineering*, *2*(5), 353–361. doi:10.1016/j.jtte.2015.08.007

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. G.P. Putmans & Sons.

Daniela, P., Calvert, T., Clark, B., & Parkhurst, G. (2019). *New Technology and Automation in Freight Transport and Handling Systems*. Academic Press.

Dantzig, G. (1949). Programming of Interdependent Activities: II Mathematical Models. *The Journal of the Operational Research Society*, *53*(11), 1275–1280.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dargis, M. (2015). Review: In 'Ex Machina', a Mogul Fashions the Droid of his dreams. *The New York Times*. https://www.nytimes.com/2015/04/10/movies/review-in-ex-machina-a-mogul-fashions-the-droid-of-his-dreams.html

Das, A., Singh, H., & Joseph, D. (2017). A longitudinal study of e-government maturity. *Information & Management*, *54*(4), 415–426. doi:10.1016/j.im.2016.09.006

Daugherty, P. R., Wilson, H. J., & Chowdhury, R. (2019). Using artificial intelligence to promote diversity. *MIT Sloan Management Review*, *60*(2), 1.

Davenport, T. H., Barth, P., & Bean, R. (2012). How Big Data Is Different. *MIT Sloan Management Review*, *54*(1), 43.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, *96*(1), 108–116.

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. doi:10.100711747-019-00696-0

Davies, J., (2016). Program good ethics into artificial intelligence. *Nature*, *538*, 291.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management*. *Information Systems*, *13*(3), 319. Advance online publication. doi:10.2307/249008

Davis, K., & Blomstrom, R. (1980). *Business and society. Concepts and policy issues* (4th ed.). McGraw Hill.

Davis, R., Vochozka, M., Vrbka, J., & Neguriţă, O. (2020). Industrial Artificial Intelligence, Smart Connected Sensors, and Big Data-driven Decision-Making Processes in Internet of Things-based Real-Time Production Logistics. *Economics*. *Management and Financial Markets*, *15*(3), 9–15.

Davoyan, A. (2020). The Impact of Artificial Intelligence on Work, Education, Mobility and Economy. In *Proceedings of the Future Technologies Conference* (pp. 291-296). Springer.

De Bruyn, A., Viswanathan, V., Beh, Y. S., Brock, J. K. U., & von Wangenheim, F. (2020). Artificial Intelligence and Marketing: Pitfalls and Opportunities. *Journal of Interactive Marketing*, *51*, 91–105. doi:10.1016/j.intmar.2020.04.007

De Caigny, A., Coussement, K., & De Bock, K. W. (2020). Leveraging fine-grained transaction data for customer life event predictions. *Decision Support Systems, 130*. . doi:10.1016/j.dss.2019.113232

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. doi:10.1016/j.ejor.2018.02.009

de Castro Neto, H., Julia, R. M. S., Paiva, E. R. F., Carvalho, A. P. L. F., Junior, A. P. S., Peres, D. R. S., . . . de Assis, J. E. (2019). Improving the AHT in telecommunication companies by automatic modeling of call center service. In *EPIA Conference on Artificial Intelligence* (pp. 96–107). Vila Real, Portugal: Springer. 10.1007/978-3-030-30244-3_9

De Keyser, A., & Köcher, S., Alkire (née Nasr), L., Verbeeck, C., & Kandampully, J. (2019). Frontline service technology infusion: Conceptual archetypes and future research directions. *Journal of Service Management*, *30*(1), 156–183. doi:10.1108/JOSM-03-2018-0082

De Keyser, A., & Lariviere, B. (2014). How technical and functional service quality drive consumer happiness: Moderating influences of channel usage. *Journal of Service Management*, *25*(1), 30–48. doi:10.1108/JOSM-04-2013-0109

de Langen, P. W. (2020). *Towards a Better Port Industry: Port Development, Management and Policy*. Taylor & Francis. doi:10.4324/9780203797501

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

de las Heras-Rosas, C., & Herrera, J. (2021). Research Trends in Open Innovation and the Role of the University. *Journal of Open Innovation*, *7*(1), 1–22. doi:10.3390/joitmc7010029

De Massis, A., Frattini, F., Kotlar, J., Petruzzelli, A. M., & Wright, M. (2016). Innovation through tradition: Lessons from innovative family businesses and directions for future research. *The Academy of Management Perspectives*, *30*(1), 93–116. doi:10.5465/amp.2015.0017

de Miranda, A. S. (2020). *Uma nova dinâmica para os portos portugueses*. PÚBLICO.

de Oliveira, W. B., & Lima, L. C. (2016). Ports and flows in the technical-scientific-informational milieu: The case of the Port of Fortaleza. *Caderno de Geografia*, *26*(46), 597–614. doi:10.5752/p.2318-2962.2016v26n.46p.597

De Quevedo, E., De la Fuente, J. M., & Delgado, J. B. (2005). Reputación Corporativa y Creación de Valor. Marco Teórico de Una Relación Circular. *Investigaciones Europeas de Dirección y Economía de la Empresa*, *11*(2), 81–97.

De Saint Laurent, C. (2018). In defence of machine learning: Debunking the myths of artificial intelligence. *Europe's Journal of Psychology*, *14*(4), 734–747. doi:10.5964/ejop.v14i4.1823 PMID:30555582

De Sousa, J. (2018). *The Challenge of Neuromarketing*. IGI Global. doi:10.4018/978-1-5225-4834-8.ch002

De Souza Nascimento, E., Ahmed, I., Oliveira, E., Palheta, M. P., Steinmacher, I., & Conte, T. (2019). *Understanding development process of machine learning systems: Challenges and solutions.* Paper presented at the 2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM). 10.1109/ESEM.2019.8870157

De Winter, J. C., & Dodou, D. (2010). Five-point Likert items: T test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, *15*(11), 2.

Delenclos, F.-X., Rasmussen, A., & Riedl, J. (2018). *To Get Smart, Ports Go Digital*. Retrieved from https://www.bcg.com/publications/2018/to-get-smart-ports-go-digital

Delenclos, F.-X., Rasmussen, A., & Riedl, J. (2018). *To Get Smart, Ports Go Digital*. BCG.

Delen, D. (2014). *Real-world data mining: Applied business analytics and decision making*. FT Press.

Delgosha, M. S., Hajiheydari, N., & Saadeatmantesh, H. (2020). Semantic structures of business analytics research: applying text mining methods. In *Information Research* (Vol. 25). University of Borås. Retrieved from http://www.informationr.net/ir/25-2/paper856.html

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, *61*(5), 637–643. doi:10.100712599-019-00595-2

Deloitte & ESPO. (2021). *Europe's ports at the crossroads of transitions A Deloitte and ESPO Study.* European Sea Ports Organisation (ESPO) and Deloitte Port Advisory.

Deloitte. (2018). *II Informe sobre la operativización de la Voz de Cliente.* Retrieved March 13, 2020, from https://www2.deloitte.com/us/en.html

Deloitte. (2021). *Smart Ports - Point of View*. Academic Press.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95. Advance online publication. doi:10.1287/isre.3.1.60

Deman, J. (2017). Star Wars is colonial fantasy: How our future imaginings are limited by our past. *The Conversation*.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Demey, J. R., Vicente-Villardon, J. L., Galindo-Villardon, M. P., & Zambrano, A. Y. (2008, December 15). (2208). Identifying molecular markers associated with classification of genotypes by external logistic biplots. *Bioinformatics (Oxford, England)*, *24*(24), 2832–2838. doi:10.1093/bioinformatics/btn552 PMID:18974073

Demirci, N. S. (2018). The Firm Specific Determinants of Fixed Capital Investments: Theories of Corporate Investment and an Applıcatıon to Industrıal Firms Listed ın BIST. *Finance & Econometrics*, 111.

Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and human-robot co-working. *Procedia Computer Science*, *158*, 688–695. doi:10.1016/j.procs.2019.09.104

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B. Methodological*, *39*(1), 1–22. doi:10.1111/j.2517-6161.1977.tb01600.x

Deng, B. (2015). Machine ethics: The robot's dilemma. *Nature*, *523*(7558), 24–26. doi:10.1038/523024a PMID:26135432

DesJardins, J. (2014). An introduction to business ethics. New York, NY: McGraw-Hill/Irwin.

Dev, C. S., Morgan, M. S., & Shoemaker, S. (1995). A positioning analysis of hotel brands: Based on travel-manager perceptions. *The Cornell Hotel and Restaurant Administration Quarterly*, *36*(6), 48–55. doi:10.1177/001088049503600617

DG MOVE. (2019). Mobility and Transport Transport in the European Union Current Trends and Issues Background Information. Brussels: Directorate-General for Mobility and Transport (DG MOVE) of the European Commission.

Dhamija, P., & Bag, S. (2020). Role of artificial intelligence in operations environment: a review and bibliometric analysis. *Business Process Management Journal, 3*(1).

Di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. (2020). Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, *121*, 283–314. doi:10.1016/j.jbusres.2020.08.019

Díaz-Bustamante, M., Carcelén, S., & Puelles, M. (2016). Image of luxury brands: A question of style and personality. *SAGE Open*, *6*(2), 1–15. doi:10.1177/2158244016644946

Diaz-Sarmiento, C., López-Lambraño, M., & Roncallo-Lafont, L. (2017). Understanding generations: a review of the concept, classification and distinguishing characteristics of baby boomers, X and millennials. *Revista Clío América*, 188-204. doi:10.21676/23897848.2440

Dickson, B. (2018). Explainable AI: Learning how AI makes decisions. *PC Mag*. Available at https://uk.pcmag.com/news/118591/learning-how-ai-makes-decisions

Dieffenbach, M. C., Gillespie, G. S., Burns, S. M., McCulloh, I. A., Ames, D. L., Dagher, M. M., Falk, E. B., & Lieberman, M. D. (2021). Neural reference groups: A synchrony-based classification approach for predicting attitudes using fNIRS. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 117–128. doi:10.1093can/nsaa115 PMID:33025001

Dietterich, T. G. (2002). Machine learning for sequential data: A review. *Lecture Notes in Computer Science, 2396*, 15–30.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology. General*, *144*(1), 114–126. doi:10.1037/xge0000033 PMID:25401381

Dignum, V., (2018). Ethics in artificial Intelligence: Introduction to the Special Issue. *Nature, Ethics and Information Technology*, *20*, 1-3.

Diwan, M. A. (2016). Internet of Things in Logistics: Towards Autonomous Logistics & Smart Logistic Entities. The International Maritime Transport & Logistics Conference (MARLOG5).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

DNV GL. (2020). *Ports: green gateways to Europe*. DNV.

Doan, T., Veira, N., Keng, B., & Ray, S. (2018). *Generating Realistic Sequences of Customer-Level Transactions for Retail Datasets*. *IEEE International Conference on Data Mining Workshops*, Singapore. 10.1109/ICDMW.2018.00122

Dodgson, M., Gann, D., & Salter, A. (2006). The role of technology in the shift towards open innovation: The case of Procter & Gamble. *R & D Management*, *36*(3), 333–346. doi:10.1111/j.1467-9310.2006.00429.x

Dolnicar, S. (2003). *Using cluster analysis for market segmentation-typical misconceptions, established methodological weaknesses and some recommendations for improvement*. Academic Press.

Domingo-Carrillo, M. A., González-Rodríguez, R., & Chávez-Miranda, E. (2020). Identifying hotel revenue management implementation drivers. *International Journal of Information Systems and Software Engineering for Big Companies*, *7*(2), 33–48.

Domingos, P. (2017). *The Master Algorithm*. Penguin Books.

Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. do i:10.1080/10618600.2017.1384734

Dooley, R. (2013). *Neuromarketing: For Coke*. It's the Real Thing.

Doree, J. (2015). *Data Science, Data Architecture-Predictive Analytics Methodology: Building a Data Map for a Greenfield Data Science Initiative*. Datasciencearchitect. Accessed and reviewed in April 2019, https://datasciencearchitect.wordpress.com/tag/togaf/

Dorrell, J., & Lee, K. (2020). The Politics of Wind: A state level analysis of political party impact on wind energy development in the United States. *Energy Research & Social Science*, *69*, 101602. doi:10.1016/j.erss.2020.101602

Doumbouya, M., Einstein, L., & Piech, C. (2020). Why AI Needs to be able to understand all the World's languages: The benefits of mobile technology are not accessible to most of the world's 700 million illiterate people. *Scientific American*. https://www.scientificamerican.com/article/why-ai-needs-to-be-able-to-understand-all-the-worlds-languages/?fbclid=IwAR3sA7mY-yDukaLdu0dC9RKCzZMEdtubdG5WoGjmk8hc4pMSY6Fdzkbt7No

DRE. (2017). Resolução do Conselho de Ministros 175/2017, 2017-11-24 - DRE. *Diário Da República n.º 227/2017, Série I de 2017-11-24*. Retrieved from https://dre.pt/home/-/dre/114248655/details/maximized

Drone deliveries take off at the port of Rotterdam. (2020, May). Retrieved from https://www.drycargomag.com/drone-deliveries-take-off-at-the-port-of-rotterdam

Droulers, O., & Roullet, B. (2010). *Neuromarketing. Le marketing revisité par les neurosciences du consommateur* [Neuromarketing. Marketing revisited by the neurosciences of the consumer?]. Dunod.

Du Pre Gauntt, J. (2008). Mobile advertising: After the growing pains. New York, NY: eMarketer.

Dubey, R., Luo, Z., Gunasekaran, A., Akter, S., Hazen, B. T., & Douglas, M. A. (2018). Big data and predictive analytics in humanitarian supply chains: Enabling visibility and coordination in the presence of swift trust. *International Journal of Logistics Management*, *29*(2), 485–512. doi:10.1108/IJLM-02-2017-0039

Ducruet, C., & Zaidi, F. (2012). Maritime constellations: A complex network approach to shipping and ports. *Maritime Policy & Management*, *39*(2), 151–168. doi:10.1080/03088839.2011.650718

Duerig, T. W., Melton, K. N., & Stöckel, D. (2013). *Engineering aspects of shape memory alloys*. Butterworth-Heinemann.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Du, S., & Xie, C. (2021). Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, *129*, 961–974. doi:10.1016/j.jbusres.2020.08.024

Dweck, C. (2019). What Having a "Growth Mindset" Actually Means. *Harvard Business Review*.

Dweck, C. S. (2017). *Summary of Carol S. Dweck's Mindset: Key Takeaways & Analysis*. CreateSpace Independent Publishing Platform.

Dwivedi, S., & Singh, V. (2013). Modeling Techniques and Applications (CIMTA) Research and reviews in question answering system. *International Conference on Computational Intelligence*. 10.1016/j.protcy.2013.12.378

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. doi:10.1016/j.ijinfomgt.2019.08.002

Ebert, C., Gallardo, G., Hernantes, J., & Serrano, N. (2016). DevOps. *IEEE Software*, *33*(3), 94–100. doi:10.1109/MS.2016.68

Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*(3), 211–218. doi:10.1007/BF02288367

Egan, J. (2011). *Relationship Marketing: Exploring Relational Strategies in Marketing*. Pearson Education.

Eissler, M. (2021). Probability that an infection like Covid-19 stops without reaching herd immunity, calculated with a stochastic agent-based model. medRxiv, 2020-12.

Ekong, U. O., Ifinedo, P., Ayo, C. K., & Ifinedo, A. (2012). E-commerce adoption in Nigerian businesses: An analysis using the technology-organization-environmental framework. In Leveraging Developing Economies with the Use of Information Technology: Trends and Tools (pp. 156-178). IGI Global.

Elsden, C., Manohar, A., Briggs, J., Harding, M., Speed, C., & Vines, J. (2018, April). Making sense of blockchain applications: A typology for HCI. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). 10.1145/3173574.3174032

Energy Market Regulatory Authority (EPDK). (2021). *Nihai Yek Listeleri*. Retrieved from http: https://www.epdk.gov.tr/Detay/Icerik/372/elektrikyekdemhttp://www.epdk.org.tr/TR/Dokuman/8692. Erişim Tarihi 07/01/2021.

English, P., & Fleischman, D. (2019). Food for thought in restaurant reviews: Lifestyle journalism or an extension of marketing in UK and Australian newspapers. *Journalism Practice*, *13*(1), 90–104. doi:10.1080/17512786.2017.1397530

ENIDH. (2021). *Escola Superior Náutica Infante D. Henrique*. Retrieved from https://www.enautica.pt/en/

Ennomotive. (2020). *Open Innovation: Definition and Types of Innovation*. Author.

Entorno turístico. (2020). *Artificial intelligence and hospitality, a good combination?* [La inteligencia artificial y el hotelería, ¿buena combinación?]. https://www.entornoturistico.com/inteligencia-artificial-hoteleria-buena-combinacion/

Erdinç, Z., & Aydınbaş, G. (2020). Panel Data Analysis on Determinants Of Renewable Energy Consumption, Journal Of Social. *Humanities and Administrative Sciences*, *6*(24), 346–358.

Erdogan, O., & Demirel, E. (2017). New Technologies in Maritime Education and Training, Turkish Experiment. *Universal Journal of Educational Research*, *5*(6), 947–952. doi:10.13189/ujer.2017.050606

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Erevelles, S., Stevenson, T. H., Srinivasan, S., & Fukawa, N. (2008). An analysis of B2B ingredient co-branding relationships. *Industrial Marketing Management*, *37*(8), 940–952. doi:10.1016/j.indmarman.2007.07.002

Eshuis, R., & Wieringa, R. (2003). Comparing Petri net and activity diagram variants for workflow modelling – a quest for reactive Petri nets. In H. Ehring, W. Reising, G. Rozenberg, & H. Weber (Eds.), Petri Net Technology for Communication-Based Systems, (vol. 2472). Springer.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. doi:10.1038/nature21056 PMID:28117445

Europa Press. (2018). *Hoteliers are betting on the Internet of Things to personalize the experience* [Los hoteleros apuestan por el Internet de las Cosas para personalizar la experiencia]. https://www.europapress.es/turismo/hoteles/noticia-hoteleros-apuestan-internet-cosas-personalizar-experiencia-20180322111219.html

European Commission. (2019), *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

European Commission. (2019). *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/futurium/en/ai-alliance-consultation

Evans, M. (2019). Build a 5-star customer experience with artificial intelligence. *Forbes*. https://www.forbes.com/sites/allbusiness/2019/02/17/customer-experience-artificial- intelligence/#1a30ebd415bd

Evers, K. (2014). *Quand la neuroscience se réveille* [When neuroscience wakes up]. Upsala Universiteit. Accessed from: https://www.crb.uu.se/digitalAssets/445/c_445294-l_1-k_k-evers-neuroethique.pdf

Evgeniou, T., Hardoon, D. R., & Ovchinnikov, A. (2020). What Happens When AI is Used to Set Grades? *Harvard Business Review*. https://hbr.org/2020/08/what-happens-when-ai-is-used-to-set-grades

Fainshtein, E., & Serova, E. (2020). Value Proposition of Network Companies Providing Restaurant Services in Russia: Analysis and Evaluation. In Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior (pp. 137-158). IGI Global.

Falk, E. B., Berkman, E. T., Whalen, D., & Lieberman, M. D. (2015). Neural Activity During Health Messaging Predicts Reductions in Smoking Above and Beyond Self-Report. *Health Psychology*. PMID:21261410

Fan, Z. P., Che, Y. J., & Chen, Z. Y. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, *74*, 90–100. doi:10.1016/j.jbusres.2017.01.010

Farah, D. (2020). *The modern MLOps blueprint*. Retrieved from https://medium.com/slalom-data-analytics/the-modern-mlops-blueprint-c8322af69d21

Farrokhi, A., Shirazi, F., Hajli, N., & Tajvidi, M. (2020). Using artificial intelligence to detect crisis related to events: Decision making in B2B by artificial intelligence. *Industrial Marketing Management*, *91*, 257–273. doi:10.1016/j.indmarman.2020.09.015

Faulhaber, A. K., Dittmer, A., Blind, F., Wächter, M. A., Timm, S., Sütfeld, L. R., Stephan, A., Pipa, G., & König, P. (2019). Human decisions in moral dilemmas are largely described by utilitarianism: Virtual car driving study provides guidelines for autonomous driving vehicles. *Science and Engineering Ethics*, *25*(2), 399–418. doi:10.100711948-018-0020-x PMID:29357047

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, *39*(11), 27–34. doi:10.1145/240455.240464

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*(2), 115–141. doi:10.1080/02699930125908

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *The Journal of Consumer Research*, *16*(3), 280–288. doi:10.1086/209214

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229–238. doi:10.1037/0022-3514.50.2.229 PMID:3701576

Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude–perception and attitude–behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology*, *51*(3), 505–514. doi:10.1037/0022-3514.51.3.505 PMID:3761146

Feng, C. M., Park, A., Pitt, L., Kietzmann, J., & Northey, G. (2020). Artificial intelligence in marketing: A bibliographic perspective. *Australasian Marketing Journal*. doi:10.1016/j.ausmj.2020.07.006

Feng, J. Z., & Jain, A. K. (2013, April). Orientation field estimation for latent fingerprint enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(4), 925–940. doi:10.1109/TPAMI.2012.155 PMID:22826508

Feng, J., Zhang, Y. Q., & Zhang, H. (2017). Improving the co-word analysis method based on semantic distance. *Scientometrics*, *111*(3), 1521–1531. doi:10.100711192-017-2286-1

Fernández, M., Puente, H., & Sequeiros, C. (2021). *Nuevas formas de comunicación a través de la narrativa en tiempo real: producción virtual, videojuegos y otras experiencias interactivas.* Congreso Internacional Comunicación y Pensamiento 2021. https://comunicacionypensamiento.org/ponencia/nuevas-formas-de-comunicacion-a-traves-de-la-narrativa-en-tiempo-real-produccion-virtual-videojuegos-y-otras-experiencias-interactivas/

Fernández, M., & Puente-Bienvenido, H. (2015). Universos fantásticos de inspiración lovecraftiana en videojuegos survival horror. Un estudio de caso de PT (Silent Hills). *Brumal. Revista de Investigación sobre lo Fantástico/Brumal. Research Journal on the Fantastic.*, *3*(1), 95–118.

Fernández-Rovira, C., Valdés, J. Á., Molleví, G., & Nicolas-Sans, R. (forthcoming). The digital transformation of business. Towards the datafication of the relationship with customers. *Technological Forecasting and Social Change, 162*, 120339.

Feuerlicht, G., & Govardhan, S. (2010). Impact of cloud computing: beyond a technology trend. *Systems Integration*.

Findasense. (2020). Estudio COVID-19. *Escenarios y Horizontes de futuro post-cuarentena en el consumo y la producción.* Retrieved November 30, 2020, from https://es.insights.findasense.com/estudios/covid-19-paper-futuro-del-consumo-y-la-produccion-66508

Fioresi, E., & Guimarães, L. (2016). *Brazilian management port: a socio-technical approach*. Academic Press.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. doi:10.1016/j.ejor.2017.11.054

Flores-Fernandez, J., & Martínez-López, F. J. (2020). Historical cycles and prospective: Our future according to our past [Ciclos históricos y prospectiva: nuestro futuro según nuestro pasado]. *Revista de Pensamiento Estratégico y Seguridad CISDE*, *5*(1), 103–121.

Flores, J., Baruca, A., & Saldivar, R. (2014). Is neuromarketing ethical? Consumers say yes. Consumers say no. Journal of Legal. *Ethical and Regulatory Issues*, *17*(2), 77–91.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1).

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Floridi, L. (2013). Distributed morality in an information society. *Science and Engineering Ethics*, *19*(3), 727–743. doi:10.100711948-012-9413-4 PMID:23197312

Fodor, O., Dell'Erba, M., Ricci, F., Spada, A., & Werthner, H. (2002). *Conceptual Normalisation of XML Data for Interoperability in Tourism. Commerce and Tourism Research Laboratory ITC-irst*. University of Trento.

Food and Agriculture Organization of the United Nations. (2017). *Food and Agriculture. Actions to advance the agenda of the 2030 Agenda and Sustainable Development Goals.* FAO. Retrieved from http://www.fao.org/3/a-i7454s.pdf

Forest, A. (1998). Cog, a Humanoid Robot, and The Question of The Image of God. *Zygon: Journal of Religion and Science*, *33*, 91-111.

Forrest, C. (2017, October 11). Google Home Mini spied on user 'thousands of times a day,' sent recordings to Google. *TechRepublic*. https://www.techrepublic.com/ article/google-home-mini-spied-on-user-thousands-of- times-a-day-sent-recordings-to-google/

Forrester Customer Experience. (2013). *Forrester's 2013 Customer Experience Predictions.* Retrieved December 3, 2020, from https://go.forrester.com/blogs/13-01-04-forresters_2013_customer_experience_predictions/

Forrester. (2020). *Improving Customer Experience By One Point Can Drive More Than A Billion Dollars In Revenue*. Retrieved December 3, 2020, from https://go.forrester.com/blogs/improving-customer-experience-by-1-point-can-drive-more-than-a-billion-dollars-in-revenue-in-2019/

Friedland, J. (2019). AI Can Help Us Live More Deliberately. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/ai-can-help-us-live-more-deliberately

Friedland, J., & Cole, B. M. (2019). From homo-economicus to homo-virtus: A system-theoretic model for raising moral self-awareness. *Journal of Business Ethics*, *155*(1), 191–205. doi:10.100710551-017-3494-6

Friedman, H. H., Herskovitz, P. J., & Pollack, S. (1993). The biasing effects of scale-checking styles on response to a Likert scale. In Proceedings of the American Statistical Association annual conference. *Survey Research Methods*, *2*, 792–795.

Frijda, N. H. (1986). The emotions. Cambridge University Press.

Fuchs, H., & Bishop, G. (1992). *Research Directions in Virtual Environments*. University of North Carolina at Chapel Hill.

Fundes Latinoamerica. (2020). *Why are FMCG companies focusing - today more than ever - on the traditional channel?* Retrieved from https://www.fundes.org/por-que-las-empresas-de-consumo-masivo-se-estan-enfocando-hoy-mas-que-nunca-en-el-canal-tradicional/

Gabriel, K. R. (1971). The biplot - graphic display of matrices with applications to principal component analysis. *Biometrika*, *58*(3), 453–467. doi:10.1093/biomet/58.3.453

Gabriel, K. R. (1995). Biplot display of multivariate categorical data, with comments on multiple correspondence analysis. In *Recent Advances in Descriptive Multivariate Analysis*. Oxford Science Publications.

Gabriel, K. R. (1998). Generalised bilinear regression. *Biometrika*, *85*(3), 689–700. doi:10.1093/biomet/85.3.689

Gabriel, K. R., & Zamir, S. (1979). And Zamir, S. (1979). Lower rank approximation of matrices by least squares with any choice of weights. *Technometrics*, *21*(4), 489–498. doi:10.1080/00401706.1979.10489819

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, *87*(3), 564–567. doi:10.1037/0033-2909.87.3.564

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Galbally, J., Haraksim, R., & Beslay, L. (2018). A study of age and ageing in fingerprint biometrics. *IEEE Transactions on Information Forensics and Security*, *14*(5), 1351–1365. doi:10.1109/TIFS.2018.2878160

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360. doi:10.1093/poq/nfp031

Galindo, P., Vicente, P., & Araya, C. (2009). Biplot Versus Coodernadas Paralelas. In Memoria del XXIII Foro de Estadística (pp. 55-59). Academic Press.

Galindo, P. (1986). Una representación simultánea: HJ Biplot. *Qüestiió. Quaderns d'Estadística, Sistemes, Informàtica i Investigació Operativ*, *10*, 13–23.

Galkin, A. (2019). *How to Calculate Price Index. Formula to Know Competitors' Impact on Your Sales in 15 Minutes*. Competera. Retrieved from https://competera.net/resources/articles/price-index-calculation

Gallagher, J. (2017, January 26). Artificial intelligence 'as good as cancer doctors'. *BBC News*. https://www.bbc. com/news/health-38717928

Gans, J., & Ryall, M. D. (2017). Value capture theory: A strategic management review. *Strategic Management Journal*, *38*(1), 17–41. doi:10.1002mj.2592

Gao, Z., Sun, D., Zhao, R., & Dong, Y. (2021, January). Ship-unloading scheduling optimization for a steel plant. *Information Sciences*, *544*, 214–226. doi:10.1016/j.ins.2020.07.029

Garbuio, M., & Lin, N. (2019). Artificial intelligence as a growth engine for health care startups: Emerging business models. *California Management Review*, *61*(2), 59–83. doi:10.1177/0008125618811931

García Haro, M. A. (2018*). La co-creación de valor a través de los social media: efecto en la imagen de Cuenca como destino turístico* (Doctoral thesis). Universidad de Castilla-La Mancha.

Garcia, A. M., & Gil-Saura, I. (2016). *Innovating in retail trade: the influence of ICT and its effects on customer satisfaction*. doi:10.5295/cdg.10556am

García-Peñalvo, F. J. (2021). Digital Transformation in the Universities: Implications of the COVID-19 Pandemic [Transformación digital en las universidades: Implicaciones de la pandemia de la COVID-19]. *Education in the Knowledge Society*, *22*, e25465. doi:10.14201/eks.25465

García-Ruiz, R., & Pérez Escoda, A. (2021). Teacher's digital competence as key for strengthening a responsible use of the Internet [La competencia digital docente como clave para fortalecer el uso responsable de Internet]. *Campus Virtuales*, *10*(1), 59–71.

Gardner, H. J., & Martin, M. A. (2007). Analyzing ordinal scales in studies of virtual environments: Likert or lump it! *Presence (Cambridge, Mass.)*, *16*(4), 439–446. doi:10.1162/pres.16.4.439

Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing Bulletin*, *2*(1), 66–70.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2016). *The need for multivocal literature reviews in software engineering: Complementing systematic literature reviews with grey literature.* Paper presented at the 20th International Conference on Evaluation and Assessment in Software Engineering. 10.1145/2915970.2916008

Gartner. (2019). *Creating a High-Impact Customer Experience Strategy*. Retrieved December 4, 2020, from https://emtemp.gcom.cloud/ngw/globalassets/en/marketing/documents/creating-a-high-impact-customer-experience-strategy-gartner-for-marketers-11-22-2019.pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gartner. (2020). *Blockchain Technology: What's Ahead? Be ready for the next phase of the blockchain revolution.* Retrieved December 3, 2020, from https://www.gartner.com/en/information-technology/insights/blockchain

Gartner. (2020). *Three Essentials for Starting and Supporting Master Data Management*. ID G00730039. Gartner Inc. Accessed and reviewed in April 2019, https://www.gartner.com/doc/reprints?id=1-24MIGFGU&ct=201119&st=sb

Gass, S. I., & Assad, A. A. (2005). *An Annotated Timeline of Operations Research*. Kluwer.

Gautier, A., Ittoo, A., & Van Cleynenbreugel, P. (2020). AI algorithms, price discrimination and collusion: A technological, economic and legal perspective. *European Journal of Law and Economics*, *50*(3), 405–435. doi:10.100710657-020-09662-6

Gebauer, H., Arzt, A., Kohtamäki, M., Lamprecht, C., Parida, V., Witell, L., & Wortmann, F. (2020). How to convert digital offerings into revenue enhancement–Conceptualizing business model dynamics through explorative case studies. *Industrial Marketing Management*, *91*, 429–441. doi:10.1016/j.indmarman.2020.10.006

Genc, R. (2010). Strategic brand management in hospitality sector: How to manage co-branding in hotels and restaurants. *Acta Universitatis Danubius Œconomica*, *6*(3), 33–46.

Genevsky, A., Yoon, C., & Knutson, B. (2017). When Brain Beats Behavior: Neuroforecasting Crowdfunding Outcomes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(36), 8625–8634. doi:10.1523/JNEUROSCI.1633-16.2017 PMID:28821681

Geraci, R. M. (2007). Robots and the Sacred in Science Fiction: Theological Implications of AI. *Zygon*, *42*, No-4.

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*(10), 2451–2471. doi:10.1162/089976600300015015 PMID:11032042

Geylani, T., Inman, J. J., & Hofstede, F. T. (2008). Image reinforcement or impairment: The effects of co-branding on attribute uncertainty. *Marketing Science*, *27*(4), 730–744. doi:10.1287/mksc.1070.0326

Gigante, M. A. (1993). Virtual reality: definitions, history and applications. *Virtual Real. Syst*., 3–14. . doi:10.1016/B978-0-12-227748-1.50009-3

Gill, R., & Singh, J. (2020). A Review of Neuromarketing Techniques and Emotion Analysis Classifiers for Visual-Emotion Mining. *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART),* 103–108. 10.1109/SMART50582.2020.9337074

Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, *115*, 106607. doi:10.1016/j.chb.2020.106607

Gill, T. (2020). Blame it on the self-driving car: How autonomous vehicles can alter consumer morality. *The Journal of Consumer Research*, *47*(2), 272–291. doi:10.1093/jcr/ucaa018

Github. (2020). *Awesome production machine learning*. Retrieved from https://github.com/EthicalML/awesome-production-machine-learning

Glaser, B. G., & Strauss, A. L. (2006). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Transactions.

Global Reporting Initiative. (2013). *Guide for the preparation of sustainability reports*. G4. Retrieved from https://www.globalreporting.org/resourcelibrary/Spanish-G4-Part-One.pdf

Gogolev, S., & Ozhegov, E. M. (2019). Comparison of Machine Learning Algorithms in Restaurant Revenue Prediction. In *International Conference on Analysis of Images, Social Networks and Texts* (pp. 27-36). Springer.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Goldsmith, J., & Burton, E. (2017). Why Teaching Artificial Intelligence Practitioners Is Important. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), 4836-4840.

Gómez Borja, M.A. (2014). *El proceso de decisión en el consumidor*. UOC.

Gómez Gutiérrez, C. (2018). United Nations Educational, Scientific and Cultural Organization (UNESCO) [El desarrollo sostenible: conceptos básicos, alcance y criterios para su evaluación]. *Tema*, III.

Gonçalves, M. J. A., Camarinha, A. P., Abreu, A. J., Teixeira, S. F., & Ferreira da Silva, A. (2020). An analysis of the most used websites in Portugal regarding accessibility web in the tourism sector. *International Journal of Information Systems and Tourism*, *5*(1), 19–28.

González-González, C. S., Violant Holz, V., Infante Moro, A., Cáceres García, L., & Guzmán Franco, M. (2021). Educational robotics in inclusive contexts: The case of the hospital classrooms [Robótica educativa en contextos inclusivos: el caso de las aulas hospitalarias]. *Educación XX1*, *24*(1), 375-403. [] doi:10.5944/educxx1.27047

González-Zamar, M. D., Abad-Segura, E., & Belmonte-Ureña, L. J. (2020). Meaningful learning in the development of digital skills. Trend analysis. *IJERI: International Journal of Educational Research and Innovation*, (14), 91–110.

González-Zamar, M. D., Abad-Segura, E., & Gallardo-Pérez, J. (2021). Ubiquitous learning in arts education and visual languages: Trend Analysis [Aprendizaje ubicuo en educación artística y lenguajes visuales: Análisis de tendencias]. *Campus Virtuales*, *10*(1), 125–139.

Goodman, E., Bamford, J., & Saynor, P. (Eds.). (2016). *Small firms and industrial districts in Italy*. Routledge.

Google Cloud. (2020). *MLOps: Continuous delivery and automation pipelines in machine learning*. Retrieved from https://cloud.google.com/solutions/machine-learning/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

Gouider, M., & Farhat, A. (2010). Mining Multi-level Frequent Itemsets under Constraints. *International Journal of Database Theory and Application*, *3*(4), 15–34.

Goulding, C. (1999). Grounded theory, ethnography and phenomenology: A comparative analysis of three qualitative strategies for marketing research. *European Journal of Marketing*, *39*(3/4), 294–308. doi:10.1108/03090560510581782

Governo do Brasil. (2021). *PSP - Porto Sem Papel*. Retrieved from https://concentrador.portosempapel.gov.br/PSP-CDP/private/comum/pages/home.xhtml

Gower, J. C., & Hand, D. (1996). *Biplots*. Chapman & Hall.

Granulo, A., Fuchs, C., & Puntoni, S. (2020). Preference for human (vs. robotic) labor is stronger in symbolic consumption contexts. *Journal of Consumer Psychology*, *31*(1), 72–80. doi:10.1002/jcpy.1181

Graziano, M. D., Renga, A., & Moccia, A. (2019). Integration of automatic identification system (AIS) data and single-channel synthetic aperture radar (SAR) images by SAR-based ship velocity estimation for maritime situational awareness. *Remote Sensing*, *11*(19), 16. doi:10.3390/rs11192196

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, *35*(6), 1241–1254. doi:10.1016/S0191-8869(02)00331-8

Greenwald, T. (2011). How smart machines like iPhone 4S are quietly changing your industry. *Forbes*. https://www.forbes.com/sites/tedgreenwald/2011/10/13/how-smart-machines-like-iphone-4s-are- quietly-changing-your-industry/#46547361598f

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *The American Psychologist*, *47*(6), 766–779. doi:10.1037/0003-066X.47.6.766 PMID:1616174

Grifoll, M., Ortego, M. I., & Egozcue, J. J. (2019). Compositional data techniques for the analysis of the container traffic share in a multi-port region. *European Transport Research Review*, *11*(1), 1. doi:10.118612544-019-0350-z

Grönroos, C. (2011). A service perspective on business relationships: The value creation, interaction and marketing interface. *Industrial Marketing Management*, *40*(2), 240–247. doi:10.1016/j.indmarman.2010.06.036

Grossman, R. P. (1997). Co-branding in advertising: Developing effective associations. *Journal of Product and Brand Management*, *6*(3), 191–201. doi:10.1108/10610429710175709

Guelfand, G. (2013). *Les études qualitatives: Fondamentaux, méthodes, analyse, techniques* [Qualitative studies: Fundamentals, methods, analysis, techniques]. EMS Editions. doi:10.3917/ems.guelf.2013.01

Guidotti, R., Monreale, A., Nanni, M., Giannotti, F., & Pedreschi, D. (2017). Clustering Individual Transactional Data for Masses of Users. In *Proceedings of the 23rd International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery. 10.1145/3097983.3098034

Guidotti, R., Rossetti, G., Pappalardo, L., Giannotti, F., & Pedreschi, D. (2019). Personalized Market Basket Prediction with Temporal Annotated Recurring Sequences. *IEEE Transactions on Knowledge and Data Engineering*, *31*(11), 2151–2163. doi:10.1109/TKDE.2018.2872587

Guillet, B. D., & Tasci, A. D. (2012). Chinese hoteliers' take on hotel co-branding in China. *Tourism Review*, *67*(4), 3–11. doi:10.1108/16605371211277777

Guimerá, A. (2015). *Mapa de la empatía (Dave Gray).* Retrieved December 5, 2020, from https://www.marketing-esencial.com/2015/03/03/mapa-de-empatia/

Güney, T. (2020). Renewable energy consumption and sustainable development in high-income countries. *International Journal of Sustainable Development and World Ecology*. Advance online publication. doi:10.1080/13504509.2020.1753124

Guo, C., Guan, Z., & Song, Y. (2012). Research on dynamic berth assignment of bulk cargo port based on ant colony algorithm. In Z. Zhang, R. Zhang, & J. Zhang (Eds.), *LISS* (pp. 235–240). Springer. doi:10.1007/978-3-642-32054-5_35

Guo, C., Guan, Z., & Song, Y. (2013). Research on bulk-cargo-port berth assignment based on priority of resource allocation. *Journal of Industrial Engineering and Management*, *6*(1), 276–288. doi:10.3926/jiem.673

Gupta, R., Khari, M., Gupta, D., & Crespo, R. G. (2020). Fingerprint image enhancement and reconstruction using the orientation and phase reconstruction. *Information Sciences*, *530*, 201–218. doi:10.1016/j.ins.2020.01.031

Guru99. (2021). *Data Science Tutorial for Beginners: What is, Basics & Process*. Accessed and reviewed in April 2019, https://www.guru99.com/data-science-tutorial.html

Gustafsod, P. E. (1998). Gender Differences in risk perception: Theoretical and methodological perspectives. *Risk Analysis*, *18*(6), 805–811. doi:10.1111/j.1539-6924.1998.tb01123.x PMID:9972583

Gutierrez, A., Boukrami, E., & Lumsden, R. (2015). Technological, organizational and environmental factors influencing managers' decision to adopt cloud computing in the UK.[. *Journal of Enterprise Information Management*, *28*(6), 788–807. doi:10.1108/JEIM-01-2015-0001

Gutmann, G., & Konagaya, A. (2019). *Predictive Simulation: Using Regression and Artificial Neural Networks to Negate Latency in Networked Interactive Virtual Reality.* arXiv preprint arXiv:1910.04703.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Haakman, M., Cruz, L., Huijgens, H., & van Deursen, A. (2020). *AI lifecycle models need to be revised. An exploratory study in fintech.* arXiv preprint arXiv:2010.02716.

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2020). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change*, *162*, 120392. doi:10.1016/j.techfore.2020.120392

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*(4), 5–14. doi:10.1177/0008125619864925

Hafkin, N. J. (2009). *E-government in africa: An overview of progress made and challenges ahead*. Paper presented at the UNDESA/UNPAN workshop on electronic/mobile government in Africa: Building Capacity in Knowledge Management through Partnership, The United Nations Economic Commission for Africa.

Hagras, H. (2018). Toward human-understandable, explainable AI. *Computer*, *51*(9), 28–36. doi:10.1109/MC.2018.3620965

Hair, J., Hult, G., Ringle, C., & Sarstedt, M. (2017). *A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.). SAGE Los Angeles.

Hakim, A., Klorfeld, S., Sela, T., Friedman, D., Shabat-Simon, M., & Levy, D. J. (2020). Machines learn neuromarketing: Improving preference prediction from self-reports using multiple EEG measures and machine learning. *International Journal of Research in Marketing*. Advance online publication. doi:10.1016/j.ijresmar.2020.10.005

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explorations*, *11*(1), 10–18. doi:10.1145/1656274.1656278

Hamilton, I. A. (2018, October 13). Why it's totally unsurprising that Amazon's recruitment AI was biased against women. *Insider*. https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10

Hansen, J. P. (2003). CAN'T MISS—Conquer Any Number Task by Making Important Statistics Simple. Part 1. Types of Variables, Mean, Median, Variance, and Standard Deviation. *Journal for Healthcare Quality*, *25*(4), 19–24. doi:10.1111/j.1945-1474.2003.tb01070.x PMID:14606209

Hao, N., Wang, H. H., & Zhou, Q. (2020). The impact of online grocery shopping on stockpile behavior in Covid-19. China agricultural economic review. *Revisión económica agrícola de China, 12*(3), 459-470. doi:10.1108/CAER-04-2020-0064

Haralambides, H. E. (2019). Gigantism in container shipping, ports and global logistics: A time-lapse into the future. *Maritime Economics & Logistics*, *21*(1), 1–60. doi:10.105741278-018-00116-0

Harari, Y. H. (2020). *The world after the coronavirus.* Retrieved from http://virtual.iuc.edu.co/pluginfile.php/3512/mod_resource/content/1/2.%20FILOSOF%C3%8DA%20MAYO%2011%20-%2015.pdf

Harkens, A. (2018). The ghost in the legal machine: Algorithmic governmentality, economy, and the practice of law. *Journal of Information Communication and Ethics in Society*, *16*(1), 16–31. doi:10.1108/JICES-09-2016-0038

Harrell, E. (2019). Neuromarketing: What you need to know. *Harvard Business Review*. Accessed from: https://hbr.org/2019/01/neuromarketing-what-you-need-to-know

Harrison-Walker, L. J. (2019). The effect of consumer emotions on outcome behaviors following service failure. *Journal of Services Marketing*, *33*(3), 285–302. doi:10.1108/JSM-04-2018-0124

Hartley, J. (2013). Some thoughts on Likert type scales. *International Journal of Clinical and Health Psychology*, *13*, 83–86.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hartley, J., & Betts, L. (2010). Four layouts and a finding: The effects of changes in the order of the verbal labels and numerical values on Likert type scales. *International Journal of Social Research Methodology*, *13*(1), 17–27. doi:10.1080/13645570802648077

Hartley, J., & Sawaya, W. (2019). Tortoise, not the hare: Digital transformation of supply chain business processes. *Business Horizons*, *62*(6), 707–715. doi:10.1016/j.bushor.2019.07.006

Hartley, P., Medlock, K. B. III, Temzelides, T., & Zhang, X. (2015). Local employment impact from competing energy sources: Shale gas versus wind generation in Texas. *Energy Econ*, *49*, 610–619. doi:10.1016/j.eneco.2015.02.023

Hart, M. C. (1996). Improving the Dissemination Of SERVQUAL By Using Magnitude Scaling. In *Total Quality Management Action* (pp. 267–271). Springer. doi:10.1007/978-94-009-1543-5_42

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, *62*(5), 565–571. doi:10.1016/j.jbusres.2008.06.016

Haşıloğlu, S. B., Baran, T., & Aydın, O. (2015). Pazarlama araştırmalarındaki potansiyel problemlere yönelik bir araştırma: Kolayda örnekleme ve sıklık ifadeli ölçek maddeleri. *Pamukkale İşletme ve Bilişim Yönetimi Dergisi*, (1), 19–28.

Hastie, T., & Tibshirani, R. (1987). Generalized Additive Models: Some Applications. *Journal of the American Statistical Association*, *82*(398), 371–386. doi:10.1080/01621459.1987.10478440

Haugland, J. (1989). *Artificial Intelligence: The Very Idea*. The Massachusetts Institute of Technology.

Heck, P. (2019). *Software engineering for machine learning applications*. Retrieved from https://fontysblogt.nl/software-engineering-for-machine-learning-applications/

Heck, P. (2020). *Testing machine learning applications*. Retrieved from https://fontysblogt.nl/testing-machine-learning-applications/

Heck, P., & Schouten, G. (2021). *Lessons learned from educating AI engineers*. Paper accepted for the 1st International Workshop on AI Engineering (WAIN).

Heeks, R. (2002). *Egovernment in Africa: Promise and practice*. Academic Press.

Hegazy, I. M. (2019). The effect of political neuromarketing 2.0 on election outcomes. *Review of Economics and Political Science*.

Heilig, L., Schwarze, S., & Voss, S. (2017). *An Analysis of Digital Transformation in the History and Future of Modern Ports*. doi:10.24251/HICSS.2017.160

Heilig, L., Lalla-Ruiz, E., & Voss, S. (2017). Digital transformation in maritime ports: Analysis and a game-theoretic framework. *NETNOMICS: Economic Research and Electronic Networking*, *18*(2-3), 227–254. Advance online publication. doi:10.100711066-017-9122-x

Heirs, S., & Manuel, M. (2021). Sustainable Maritime Career Development: A case for Maritime Education and Training (MET) at the Secondary Level. *TransNav: International Journal on Marine Navigation & Safety of Sea Transportation*, *15*(1), 91–99. doi:10.12716/1001.15.01.08

Helmig, B., Huber, J.-A., & Leeflang, P. S. (2008). Co-branding: The state of the art. *Schmalenbach Business Review*, *60*(4), 359–377. doi:10.1007/BF03396775

Hendarman, A. F., & Tjakraatmadja, J. H. (2012). Relationship among soft skills, hard skills, and innovativeness of knowledge workers in the knowledge economy era. *Procedia: Social and Behavioral Sciences*, *52*, 35–44. doi:10.1016/j.sbspro.2012.09.439

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Henn na Hotel. (n.d.). https://www.h-n-h.jp/en/concept

Hermann, K. M., Kocisky, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., & Blunsom, R. (2015). Teaching machines to read and comprehend. In *28th International Proceedings on Advances in Neural Information Processing Systems* (pp. 1693–1701). MIT Press.

Herodotou, H., Aslam, S., Holm, H., & Theodossiou, S. (2020). Big Maritime. *Data Management*, 313–334. doi:10.1007/978-3-030-50892-0_19

Hertzum, M. (2018). Three contexts for evaluating organizational usability. *Journal of Usability Studies Archive*, *14*, 35–47.

Hertzum, M. (2021). *Organizational Implementation: The Design in Use of Information Systems*. Morgan & Claypool Publishers.

Hesenius, M., Schwenzfeier, N., Meyer, O., Koop, W., & Gruhn, V. (2019). *Towards a software engineering process for developing data-driven applications.* Paper presented at the 2019 IEEE/ACM 7th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE). 10.1109/RAISE.2019.00014

Hicks, S. C., & Irizarry, R. A. (2018). A Guide to Teaching Data Science. *The American Statistician*, *72*(4), 382–391. doi:10.1080/00031305.2017.1356747 PMID:31105314

Hinchliffe, E. (2017, February 22), IBM's Watson will help diagnose heart disease when doctors may have missed the signs. *Mashable*. https://mashable.com/2017/02/22/ibm-watson-clinical-imaging-cardiology/

Hirsch, J. (2020). *An fNIRS Approach to Two-person Neuroscience*. Biophotonics Congress: Biomedical Optics 2020 (Translational, Microscopy, OCT, OTS, BRAIN) (2020), Paper BM4C.1, BM4C.1. 10.1364/BRAIN.2020.BM4C.1

Hirsch, J., Zhanga, X., Noah, J. A., & Ono, Y. (2017). Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *NeuroImage*, *157*, 314–330. doi:10.1016/j.neuroimage.2017.06.018 PMID:28619652

Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the internet of things: An assemblage theory approach. *The Journal of Consumer Research*, *44*(6), 1178–1204. doi:10.1093/jcr/ucx105

Hofmans, J., Theuns, P., Baekelandt, S., Mairesse, O., Schillewaert, N., & Cools, W. (2007). Bias and changes in perceived intensity of verbal qualifiers effected by scale orientation. *Survey Research Methods*, *1*(2), 97–108.

Ho, L. C. J., & Taylor, M. E. (2007). An empirical analysis of triple bottom-line reporting and its determinants: Evidence from the United States and Japan. *Journal of International Financial Management & Accounting*, *18*(2), 123–150. doi:10.1111/j.1467-646X.2007.01010.x

Holme, R., & Watts, P. (2001). Making Good Business Sense. *Journal of Corporate Citizenship*, *2001*(2), 17–20. doi:10.9774/GLEAF.4700.2001.su.00005

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1), 212–222. doi:10.1016/j.neuroimage.2012.06.028 PMID:22732563

Holzer, S., & Andruet, R. (1995). *A multimedia workshop learning environment for statics*. doi:10.1109/FIE.1995.483056

Hong, J. W. (2020). Why Is Artificial Intelligence Blamed More? Analysis of Faulting Artificial Intelligence for Self-Driving Car Accidents in Experimental Settings. *International Journal of Human-Computer Interaction*, *36*(18), 1768–1774. doi:10.1080/10447318.2020.1785693

Hong, W., & Thong, J. (2013). Internet privacy concerns: An integrated conceptualisation and four empirical studies. *Management Information Systems Quarterly*, *37*(1), 275–298. doi:10.25300/MISQ/2013/37.1.12

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hossain, F. T., Hossain, M. I., & Nawshin, S. (2017). *Machine learning based class level prediction of restaurant reviews. In 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE.

Hossain, M., Leminen, S., & Westerlund, M. (2019). A systematic review of living lab literature. *Journal of Cleaner Production*, *213*, 976–988. doi:10.1016/j.jclepro.2018.12.257

Howard, M. (2020, March). Water drone christened in the port of Hamburg. *Maritime Logistics Professional*.

Howard, M., Steensma, H. K., Lyles, M., & Dhanaraj, C. (2016). Learning to collaborate through collaboration: How allying with expert firms influences collaborative innovation within novice firms. *Strategic Management Journal, 37*(10), 2092–2103. doi:10.1002mj.2424

Hsu, L., & Chen, Y.-J. (2020). Neuromarketing, subliminal advertising, and hotel selection: An EEG study. *Australasian Marketing Journal*, *28*(4), 200–208. Advance online publication. doi:10.1016/j.ausmj.2020.04.009

Huang, B., Laporte, S., Sénécal, S., & Sobol, K. (2021), Partner or servant? The effect of relationship role on consumer interaction with artificial intelligence. *AMA Winter Academic Conference Proceedings*, *32*, 382-384.

Huang, M., Rust, R.T., (2018). Artificial Intelligence in Service. *Journal of Service Research*, *2*(2), 155-172.

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, *21*(2), 155–172. doi:10.1177/1094670517752459

Hulten, G. (2019). *Building intelligent systems*. Apress. doi:10.1007/978-1-4842-3933-9

Hurst, J. (2015). Assessing kindship through scent – the mouse model. *Chemical Senses*, *40*(3), 211–297. doi:10.1093/chemse/bju073

Hustić, I., & Gregurec, I. (2015). *The influence of price on customer's purchase decision*. Academic Press.

Hwang, H., & Park, M. H. (2020). The Threat of AI and Our Response: The AI Charter of Ethics in South Korea. *Asian Journal of Innovation & Policy*, *9*(1), 56–78.

Hwang, J. S. (2016). The fourth industrial revolution (industry 4.0): Intelligent manufacturing. *SMT Magazine*, *3*, 616–630.

iabSpain. (2018). *Estudio anual de eCommerce España 2018*. Retrieved December 6, 2020, from https://iabspain.es/wp-content/uploads/estudio-ecommerce-iab-2018_vcorta.pdf

Iacovou, C. L., Benbasat, I., & Dexter, A. S. (1995). Electronic data interchange and small organizations: Adoption and impact of technology. *Management Information Systems Quarterly*, *9*(4), 465–485. doi:10.2307/249629

IBM. (2015a). *Modelling the Entity data architecture*. IBM.

IBM. (2019). *AI and Data Management-Delivering Data-Driven Business Transformation and Operational Efficiencies*. IBM USA.

Ibrahim, A. M. (2001). Differential responding to positive and negative items: The case of a negative item in a questionnaire for course and faculty evaluation. *Psychological Reports*, *88*(2), 497–500. doi:10.2466/pr0.2001.88.2.497 PMID:11351897

IEEE. (2019). *Ethically Aligned Design*. Institute of Electrical and Electronic Engineers. https://ethicsinaction.ieee.org/

Ifinedo, P. (2012). Factors influencing e-government maturity in transition economies and developing countries: A longitudinal perspective. *SIGMIS Database*, *42*(4), 98–116. doi:10.1145/2096140.2096147

IMD. (2015). *IMD business school and Cisco join forces on digital business transformation*. IMD.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

INEI. (2014). *Results of the Micro and Small Enterprise Survey 2013.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1139/libro.pdf

INEI. (2017). *National Census of Food Markets 2016.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1448/libro.pdf

INEI. (2018). *Technical report Estadísticas de las Tecnologías de Información y Comunicación en los Hogares.* Retrieved from https://www.inei.gob.pe/media/MenuRecursivo/boletines/04-informe-tecnico-n04_tecnologias-de-informacion-jul-ago-set2018.pdf

Infante Moro, J. C. (2017). *User perception to improve the use of Social Networks as a communication channel in the hotel sector* [Percepción de los usuarios para la mejora del uso de las Redes Sociales como canal de comunicación en el sector hotelero] (Doctoral Thesis). University of Huelva.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021a). Factors that influence the adoption of the Internet of Things in the hotel sector [Factores que influyen en la adopción del Internet de las Cosas en el sector hotelero]. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, (E41), 370-383.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020a). Key factors in the implementation of Cloud Computing as a service and communication tool in universities. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 631-636). 10.1145/3434780.3436698

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020b). Motivational factors in the insertion of digital skills in teaching. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 365-370). 10.1145/3434780.3436631

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020c). Motivational factors that justify the implementation of the internet of things as a security system in the hotel sector [Factores motivacionales que justifican la implementación del Internet de las Cosas como sistema de seguridad en el sector hotelero]. *Revista de Pensamiento Estratégico y Seguridad CISDE*, *5*(2), 81–91.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020d). The employment possibilities of the internet of things in the hotel sector and its training needs [Las posibilidades de empleo del Internet de las Cosas en el sector hotelero y sus necesidades formativas]. *Education In The Knowledge Society*, *21*, 14. doi:10.14201/eks.222643

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021b). The acquisition of ICT skills at the university level: The case of the Faculty of Business Studies and Tourism of the University of Huelva [La adquisición de competencias TIC en el ámbito universitario: El caso de la Facultad de Ciencias Empresariales y Turismo de la Universidad de Huelva]. *Revista de Medios y Educacion*, (60), 29–58.

Infante-Moro, A., Infante-Moro, J. C., & Gallardo-Pérez, J. (2021c). Key Factors in the Implementation of the Internet of Things in the Hotel Sector. *Applied Sciences (Basel, Switzerland)*, *11*(7), 2924. doi:10.3390/app11072924

Information Desk. (2021). *Ritwik Ghatak's War between 'Man and Machine*. Get Bengal, Thinking Positive.

Inkinen, T., Helminen, R., & Saarikoski, J. (2019). Port Digitalization with Open Data: Challenges, Opportunities, and Integrations. *Journal of Open Innovation*, *5*(2), 30. doi:10.3390/joitmc5020030

Inselberg, A. (1992). *The plane R2 with coordinate parallel. Computer Science and Applied Mathematics Departments. Tel Aviv*. University.

Institute for Ethical AI & Machine Learning. (2020). *Machine learning maturity model v1.0*. Retrieved from https://ethical.institute/mlmm.html

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Institute of Chartered Accountants in England and Wales (ICAEW). (2004). *Sustainability: The Role of Accountants.* Sustainable Business Initiative Report.

Instıtute, T. S. (TUIK). (2021). Retrieved from http: https://data.tuik.gov.tr/

International Council for Local Environmental Initiatives (ICLEI). (2021). *Climate Protection*. Retrieved from http: https://iclei-europe.org/our-members/

International Energy Agency (IEA). (2021). *Data and statistics.* Retrieved from http: https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20supply&indicator=TPESbySource

Internet World Sats. (2020). *Internet World Sats*. Retrieved from https://www.internetworldstats.com/stats2.htm

Intersoft Consulting. (2020). *General Data Protection Regulation GDPR*. Retrieved December 6, 2020, from https://gdpr-info.eu/

Ioannidou. (2013). Adapting Superhero Comics for the Big Screen: Subculture for the Masses. *Adaptation*, *6*(2), 230–238. doi:10.1093/adaptation/apt004

Ishiguro, K. (2021). *Klara and the Sun*. Faber and Faber.

Ishikawa, F., & Yoshioka, N. (2019). *How do engineers perceive difficulties in engineering of machine-learning systems? Questionnaire survey.* Paper presented at the 2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SER&IP). 10.1109/CESSER-IP.2019.00009

İşlek, İ., & Gunduz Oguducu, S. (2017). A Decision Support System for Demand Forecasting based on Classifier Ensemble. *Federated Conference on Computer Science and Information Systems (ACSIS)*, 13, 35–41. 10.15439/2017F224

ISO. (2020). *ISO/IEC 25010*. Retrieved December 18, 2020, from https://iso25000.com/index.php/normas-iso-25000/iso-25010

itUser. (2020). *La experiencia de cliente y el aumento de ingresos, una relación que confirman el 82% de las empresas.* Retrieved December 18, 2020, from https://www.ituser.es/estrategias-digitales/2020/06/la-experiencia-de-cliente-y-el-aumento-de-ingresos-una-relacion-que-confirman-el-82-de-las-empresas

Ivanov, O. A., Ivanova, V. V., & Saltan, A. A. (2018). Likert scale questionnaires as an educational tool in teaching discrete mathematics. *International Journal of Mathematical Education in Science and Technology*, *49*(7), 1110–1118. doi:10.1080/0020739X.2017.1423121

Iwashita, M. (2019). Transitional Method for Identifying Improvements in Video Distribution Services. *International Journal of Networked and Distributed Computing*, *7*(4), 141. doi:10.2991/ijndc.k.190911.001

Izard, C. E. (1977). *Human emotions*. Plenum Press.

Izo. (2019a). *5 Tendencias en Experiencia de Cliente en 2019*. Retrieved December 18, 2020, from https://izo.es/tendencias-experiencia-de-cliente-2019/

Izo. (2019b). *¿Cómo Conseguir Empleados Felices y Comprometidos?* https://izo.es/como-conseguir-empleados-felices-y-comprometidos/

Jabrouni, H., Kamsu-Foguem, B., Geneste, L., & Vaysse, C. (2011). Continuous improvement through knowledge-guided analysis in experience feedback. *Eng. Appl. of AI*, *24*(8), 1419–1431. doi:10.1016/j.engappai.2011.02.015

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jacks, D. S., & Pendakur, K. (2010). Global Trade and The Maritime Transport Revolution. *The Review of Economics and Statistics*, *92*(4), 745–755. doi:10.1162/REST_a_00026

Jackson, L.A., Witt, E., Games, A., Fitzgerald, H., Von Eye, A. & Zao, Y. (2011). Information Technological use and Creativity: Findings from The Children and Technological Project. *Computers in Human Behaviour, 28*(2), 370-376.

Jacoby, J., & Matell, M. S. (1971). Three-point Likert scales are good enough. *JMR, Journal of Marketing Research*, *8*(4), 495–500. doi:10.1177/002224377100800414

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323. doi:10.1145/331499.331504

Jamieson, S. (2004). Likert scales: How to (ab) use them. *Medical Education*, *38*(12), 1217–1218. doi:10.1111/j.1365-2929.2004.02012.x PMID:15566531

Jamil, M. G., & Bhuiyan, Z. (2021). Deep learning elements in maritime simulation programs: A pedagogical exploration of learner experiences. *International Journal of Educational Technology in Higher Education*, *18*(1), 1–22. doi:10.118641239-021-00255-0

Jannach, D., Zanker, M., & Fuchs, M. (2014). Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. *Information Technology & Tourism*, *14*(2), 119–149. doi:10.100740558-014-0010-z

Jansen, A. (2006). *Assessing e-government progress– Why and what.* https://www.uio.no/studier/emner/jus/afin/FINF4001/h05/undervisningsmateriale/AJJ-nokobit2005.pdf

Jarek, K., & Mazurek, G. (2019). Marketing and Artificial Intelligence. *Central European Business Review*, *8*(2), 46–55. doi:10.18267/j.cebr.213

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision-making. *Business Horizons*, *61*(4), 577–586. doi:10.1016/j.bushor.2018.03.007

Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality & Quantity, 41*(1), 37–54. https://link.springer.com/article/10.1007/s11135-005-5427-1

Jaszkiewicz, J., & Sowiñski, R. (1999). The 'Light Beam Search' approach - an overview of methodology and applications. *European Journal of Operational Research*, *113*(2), 300–314. doi:10.1016/S0377-2217(98)00218-5

Jean, W., & Yazdanifard, R. (2015). The Review of How Sales Promotion Change the Consumer's Perception and Their Purchasing Behavior of a Product. *Global Journal of Management and Business Research*, *15*, 33–37.

Jenkins, H. (2007). *Transmedia Storytelling 101.* http://henryjenkins.org/2007/03/transmedia_storytelling_101.html

Jenkins, H. (2011). *Transmedia 202: Further Reflections.* http://henryjenkins.org/2011/08/defining_transmedia_further_re.html

Jenkins, H. (2013). *"Wish You Were Here": Imaginative Mobilities and Disembodied Intimacy in Postcards (Part Two).* http://henryjenkins.org/2013/02/what-transmedia-producers-need-toknow-about-comics-an-interview-with-tyler-weaver-part-two.html

Jenkins, H. (2006). *Convergence Culture*. New York University Press.

Jensen, Ø., & Hansen, K. V. (2007). Consumer values among restaurant customers. *International Journal of Hospitality Management*, *26*(3), 603–622. doi:10.1016/j.ijhm.2006.05.004

Jerz, D. G. (2002). R.U.R. (Rossum's Universal Robots). Dover.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jing, L., Marlow, P. B., & Hui, W. (2008). An analysis of freight rate volatility in dry bulk shipping markets. *Maritime Policy & Management*, *35*(3), 237–251. doi:10.1080/03088830802079987

Jin, N., Lee, S., & Huffman, L. (2012). Impact of restaurant experience on brand image and customer loyalty: Moderating role of dining motivation. *Journal of Travel & Tourism Marketing*, *29*(6), 532–551. doi:10.1080/10548408.2012.701552

Johnson, J. M., Bristow, D. N., & Schneider, K. C. (2004). Did you not understand the question or not? An investigation of negatively worded questions in survey research. *Journal of Applied Business Research*, *20*(1), 75–86.

Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion, and Parties*, *15*(2), 237–264. doi:10.1080/13689880500178849

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In *Handbook of Emotions* (2nd ed., pp. 220–235). Guilford Press.

Jones, C., Alderete, M. V., & Motta, J. (2013). E-commerce adoption in Micro, Small and Medium-sized commercial and service enterprises in Cordoba. *Cuadernos de Administración, 29*(50), 164-175. Retrieved from https://www.redalyc.org/pdf/2250/225029797006.pdf

Jones, V. K. (2018). Voice-activated change: Marketing in the age of artificial intelligence and virtual assistants. *Journal of Brand Strategy*, *7*(3), 233–245.

Jose, P. C., & Fernandes, A. P. (2020). European Seaports Information Systems. The Impacts of Directive 2010/65/EU. *Economics and Culture*, *17*(2), 38–49. doi:10.2478/jec-2020-0019

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *Current Journal of Applied Science and Technology*, 396-403.

Joy, B. (2000). *Why the future doesn't need us*. Wired.

Józsa, K., & Morgan, G. A. (2017). Reversed items in Likert scales: Filtering out invalid responders. *Journal of Psychological and Educational Research*, *25*(1), 7–25.

June Kim, S. (2021). How can higher maritime education lead shipping growth? Korea's experience, 1948–1982. *International Journal of Maritime History*, *33*(1), 90–117. doi:10.1177/0843871420974062

Jung, T., & Dieck, T. (2018). *Augmented Reality and Virtual Reality. Empowering Human, Place and Business*. Springer. doi:10.1007/978-3-319-64027-3

Junta de Andalucía. (2020). *Guide Using ICTs to reach new market niches and consumers.* Retrieved from www.formacion.andaluciaesdigital.es: https://www.formacion.andaluciaesdigital.es/catalogo-cursos/-/acciones/ficha/7809

Jurney, R. (2017). *Agile data science 2.0: Building full-stack data analytics applications with Spark*. O'Reilly Media.

Jussila, J., Huhtamäki, J., Kärkkäinen, H., Aho, T., Kortelainen, S., & Tebest, T. (2014). New era of business analytics - Making sense of business ecosystems. *MINDTREK 2014 - Proceedings of the 18th International Academic MindTrek Conference*, 266–268. 10.1145/2676467.2676517

Kadircan, H. K. (2019). Medical Ethics: Consideration on Artificial Intelligence. *Journal of Clinical Neuroscience*.

Kak, S. (2006). Artificial and biological intelligence. *ACM Ubiquity*, *6*(42), 1–20.

Kaloudi, N., & Li, J. (2020). The AI-based cyber threat landscape: A survey. *ACM Computing Surveys*, *53*(1), 1–34. doi:10.1145/3372823

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kalouptsidi, M. (2014, February). Time to build and fluctuations in bulk shipping. *The American Economic Review*, *104*(2), 564–608. doi:10.1257/aer.104.2.564

Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, *15*(1), 276. doi:10.1186/1471-2105-15-276 PMID:25123979

Kang, H., Le, M., & Tao, S. (2016). *Container and microservice driven design for cloud infrastructure devops.* Paper presented at the 2016 IEEE International Conference on Cloud Engineering (IC2E). 10.1109/IC2E.2016.26

Kang, K., Zhang, W.-c., Guo, L.-y., & Ma, T. (2012, September). Research on ship routing and deployment mode for a bulk. In *2012 international conference on management science & engineering 19th annual conference proceedings.* Dallas, TX: IEEE. 10.1109/ICMSE.2012.6414421

Kantar. (2020). *A new world: Changes and prospects for shopping channels.* Retrieved from http://mkt.kantarworldpanel.com: http://mkt.kantarworldpanel.com/global/LATAM/Retail/Retail_COVID-19_v3.pdf

Kapferer, J.-N. (1998). Why are we seduced by luxury brands? *Journal of Brand Management*, *6*(1), 44–49. doi:10.1057/bm.1998.43

Kapferer, J.-N. (2012). *The New Strategic Brand Management: Advanced Insights and Strategic Thinking* (5th ed.). Kogan Page.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. doi:10.1016/j.bushor.2018.08.004

Karhade, P. P., & Dong, J. Q. (2021). Innovation Outcomes of Digitally Enabled Collaborative Problemistic Search Capability. *Management Information Systems Quarterly*, *45*(2), 693–717. doi:10.25300/MISQ/2021/12202

Karlis, T. (2018, January). Maritime law issues related to the operation of unmanned autonomous cargo ships. *WMU Journal of Maritime Affairs*, *17*(1), 1–10. doi:10.100713437-018-0135-6

Karlsson, L. C., Soveri, A., Lewandowsky, S., Karlsson, L., Karlsson, H., Nolvi, S., Karuki, M., Lindfelt, M., & Antfolk, J. (2021). Fearing the disease or the vaccine: The case of COVID-19. *Personality and Individual Differences*, *172*, 110590. doi:10.1016/j.paid.2020.110590 PMID:33518869

Karnowski, J., & Fadely, R. (2016). *How AI careers fit into the data landscape*. Retrieved from https://blog.insightdatascience.com/how-emerging-ai-roles-fit-in-the-data-landscape-d4cd922c389b

Kavallieratos, G., Diamantopoulou, V., & Katsikas, S. K. (2020). Shipping 4.0: Security Requirements for the Cyber-Enabled Ship. *IEEE Transactions on Industrial Informatics*, *16*(10), 6617–6625. doi:10.1109/TII.2020.2976840

Kaviya, K., Roshini, C., Vaidhehi, V., & Sweetlin, J. D. (2017). Sentiment analysis for restaurant rating. In *2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)* (pp. 140-145). IEEE.

Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010). *Mastering the information age: Solving problems with visual analytics*. VisMaster.

Keller, K. L., & Richey, K. (2006). The importance of corporate brand personality traits to a successful 21st century business. *Journal of Brand Management*, *14*(1-2), 74–81. doi:10.1057/palgrave.bm.2550055

Kennedy, M., Campbell, I., & Chyung, Y. (2018). *Evidence-Based Survey Design: The Use of Ascending or Descending Order of Likert Response Options*. Academic Press.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kent, S. (2001). *The Ultimate History of Video Games: From Pong to Pokemon—The Story Behind the Craze That Touched Our Lives and Changed the World*. Crown.

Kerr, I., & Earle, J. (2013). Prediction, preemption, presumption: How big data threatens big picture privacy. *Stanford Law Review Online*, *66*, 65–72.

Kery, M. B., Radensky, M., Arya, M., John, B. E., & Myers, B. A. (2018). *The story in the notebook: Exploratory data science using a literate programming tool.* Paper presented at the CHI Conference on Human Factors in Computing Systems. 10.1145/3173574.3173748

Kęsek, M., Bogacz, P., & Migza, M. (2019). The application of Lean Management and Six Sigma tools in global mining enterprises. *IOP Conference Series. Earth and Environmental Science*, *214*, 12090. doi:10.1088/1755-1315/214/1/012090

Kevin, W. J. (2000). *Co-founder of Sun Microsystem-expressed in a 2000 Wired Magazine article*. Academic Press.

Khan, H., Khan, I., Oanh, L. T. K., & Lin, Z. (2020). The Dynamic Interrelationship of Environmental Factors and Foreign Direct Investment: Dynamic Panel Data Analysis and New Evidence from the Globe. *Mathematical Problems in Engineering*, 1–12.

Khan, M. A., Khan, T. M., Bailey, D. G., & Kong, Y. (2016). A spatial domain scar removal strategy for fingerprint image enhancement. *Pattern Recognition*, *60*, 258–274. doi:10.1016/j.patcog.2016.05.015

Khan, Z., Malik, M. Y., Latif, K., & Jiao, Z. (2020). Heterogeneous effect of eco-innovation and human capital on renewable & non-renewable energy consumption: Disaggregate analysis for G-7 countries. *Energy*, *209*, 118405. doi:10.1016/j.energy.2020.118405

Kim, H. J., Pan, G., & Pan, S. L. (2007). Managing it-enabled transformation in the public sector: A case study on e-government in south korea. *Government Information Quarterly*, *24*(2), 338–352. doi:10.1016/j.giq.2006.09.007

Kim, H. W., Chan, H. C., & Gupta, S. (2007). Value-based Adoption of Mobile Internet: An empirical investigation. *Decision Support Systems*, *43*(1), 111–126. doi:10.1016/j.dss.2005.05.009

Kim, K., & Kim, K. (1999). Routing straddle carriers for the loading operation of containers using a beam search algorithm. Elsevier. *Computers & Industrial Engineering*, *36*(1), 109–136. doi:10.1016/S0360-8352(99)00005-4

Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2017). Data scientists in software teams: State of the art and challenges. *IEEE Transactions on Software Engineering*, *44*(11), 1024–1038. doi:10.1109/TSE.2017.2754374

Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction. *International Journal of Contemporary Hospitality Management*, *29*(2), 784–802. doi:10.1108/IJCHM-11-2015-0627

Kinder, M. (1991). *Playing with Power in Movies, Television and Video Games*. University of California Press.

Kinder, M. (2014). *Transmedia Frictions: The Digital, the Arts, and the Humanities*. University of California Press.

Kingma, D. P., & Ba, J. (2015). Adam: {A} Method for Stochastic Optimization. In Y. Bengio & Y. LeCun (Eds.), *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015*, *Conference Track Proceedings*. https://arxiv.org/abs/1412.6980

Kirk, C. P., & Rifkin, L. S. (2020, September). I'll trade you diamonds for toilet paper: Consumer reacts, copes, and adapts behaviors in the COVID-19 pandemic. *Journal of Business Research*, *117*, 124–131. doi:10.1016/j.jbusres.2020.05.028 PMID:32834208

Kirk, M. (2017). *Thoughtful machine learning with Python: A test-driven approach*. O'Reilly Media.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Klass, M. (1983). The Artificial Alien: Transformations of the Robot in Science Fiction. *Annals of the American Academy of Political and Social Science*, 171-79.

Kleinginna, P.R., & Kleinginna, A.M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, *5*(4), 345–379.

KLM. (2016, March 29). *Spencer robot completed tests guiding KLM passengers at Schiphol*. https://news.klm.com/spencer-robot-completed-tests-guiding-klm-passengers-at-schiphol/

Klmcare. (n.d.). *Meet KLM Care-E*. http://klmcare-e-entry.com

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, *39*(2), 121–123. doi:10.1097/00006199-199003000-00019 PMID:2315066

Knowledge Management. (2021). *Experience-Based Knowledge*. Knowledge management, Group Work, UCPori. Retrieved from https://knowledgemanagement5.wordpress.com/experience-based-knowledge/

Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, *53*(1), 147–156. doi:10.1016/j.neuron.2006.11.010 PMID:17196537

Kodratoff, Y. (1999, July). Comparing machine learning and knowledge discovery in databases: An application to knowledge discovery in texts. In *Advanced Course on Artificial Intelligence* (pp. 1–21). Springer.

Kohli, A. K., Jaworski, B. J., & Shabshab, N. (2019). Customer centricity: a multi-year journey. In *Handbook on Customer Centricity*. Edward Elgar Publishing.

Kolarik, T., & Rudorfer, G. (1994). Time series forecasting using neural networks. *Proceedings of the International Conference on APL : The Language and Its Applications, APL 1994*, 86–94. 10.1145/190271.190290

Kolb, D., & Kolb, A. (2017). *The Experiential Educator: Principles and Practices of Experiential Learning*. Academic Press.

Koller, B. (2020, Sep 28). *12 Factors of reproducible Machine Learning in production*. Retrieved from https://blog.maiot.io/12-factors-of-ml-in-production/

Korinek, A., & Stiglitz, J. (2021). *Artificial Intelligence Could Mean Large Increase in Prosperity - But only for a Privileged Few*. https://www.ineteconomics.org/perspectives/blog/artificial-intelligence-could-mean-technological-advancement-but-only-for-a-privileged-few?fbclid=IwAR0qxVBh0NYHEJLVJCoC3bynkESnLFgfn8QFZcIE63kGu06TIjMEbWDHadM

Kotler, P., Armstrong, G., Cámara, D., & Cruz, I. (2004). Principios de marketing (10a edición). Madrid: Prentice Hall.

Kraus, S. J. (1995). Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature. *Personality and Social Psychology Bulletin*, *21*(1), 58–75. doi:10.1177/0146167295211007

Krensky, P., den Hamer, P., Brethenoux, E., Hare, J., Idoine, C., Linden, A., Sicular, S., & Choudhary, F. (2020). *Magic quadrant for data science and machine learning platforms*. Academic Press.

Kreps, S., Prasad, S., Brownstein, J. S., Hswen, Y., Garibaldi, B. T., Zhang, B., & Kriner, D. L. (2020). Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA Network Open*, *3*(10), e2025594–e2025594. doi:10.1001/jamanetworkopen.2020.25594 PMID:33079199

Kretschmann, L., Zacharias, M., Klöver, S., & Hensel, T. (2020). *Machine Learning in Maritime Logistics*. Fraunhofer CML.

Kriens, P., & Verbelen, T. (2019). *Software engineering practices for machine learning*. arXiv preprint arXiv:1906.10366.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Krish, R. P., Fierrez, J., Ramos, D., Alonso-Fernandez, F., & Bigun, J. (2019). Improving automated latent fingerprint identification using extended minutia types. *Information Fusion*, *50*, 9–19. doi:10.1016/j.inffus.2018.10.001

Krueger, M. (1991). *Realidad artificial 2*. Addison-Wesley Profesional.

Krueger, M. W., Gionfriddo, T., & Hinrichsen, K. (1985). Videoplace an artificial reality. *Proceedings of the ACM SIGCHI Bulletin*, 16. 10.1145/317456.317463

Kühn, A., Joppen, R., Reinhart, F., Röltgen, D., von Enzberg, S., & Dumitrescu, R. (2018). Analytics Canvas–A Framework for the design and specification of data analytics projects. *Procedia CIRP*, *70*, 162–167. doi:10.1016/j.procir.2018.02.031

Kula, V., & Tatoglu, E. (2003). An exploratory study of Internet adoption by SMEs in an emerging market economy.[. *European Business Review*, *15*(5), 324–333. doi:10.1108/09555340310493045

Kumar, S., Yadava, M., & Roy, P. P. (2019). Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Information Fusion*, *52*, 41–52. doi:10.1016/j.inffus.2018.11.001

Kumar, V., & Leone, R. P. (1988). Measuring the Effect of Retail Store Promotions on Brand and Store Substitution. *JMR, Journal of Marketing Research*, *25*(2), 178–185. doi:10.1177/002224378802500206

Kunstelj, M., & Vintar, M. (2004). Evaluating the progress of e-government development: A critical analysis. *Info. Pol.,* *9*(3,4), 131-148.

Kurzweil, R. (1999). *The Age of Spiritual Machines*. Viking Adult.

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking Books.

La Torre, M., Botes, V. L., Dumay, J., & Odendaal, E. (2019). Protecting a new achilles heel: The role of auditors within the practice of data protection. *Managerial Auditing Journal*. Advance online publication. doi:10.1108/MAJ-03-2018-1836

La Torre, M., Botes, V. L., Dumay, J., Rea, M. A., & Odendaal, E. (2018a). The fall and rise of intellectual capital accounting: New prospects from the big data revolution. *Meditari Accountancy Research*, *26*(3), 381–399. doi:10.1108/MEDAR-05-2018-0344

La Torre, M., Dumay, J., & Rea, M. A. (2018b). Breaching intellectual capital: Critical reflections on big data security. *Meditari Accountancy Research*, *26*(3), 463–482. doi:10.1108/MEDAR-06-2017-0154

Laforet, S. (2011). A framework of organisational innovation and outcomes in SMEs. *International Journal of Entrepreneurial Behaviour & Research*, *17*(4), 380–408. doi:10.1108/13552551111139638

Laforet, S., & Tann, J. (2006). Innovative characteristics of small manufacturing firms. *Journal of Small Business and Enterprise Development*, *13*(3), 363–380. doi:10.1108/14626000610680253

Lambrou, M., Watanabe, D., & Iida, J. (2019, November). Shipping digitalization management: Conceptualization, typology, and antecedents. *Journal of Shipping and Trade*, *4*(11), 11. Advance online publication. doi:10.118641072-019-0052-7

Lanctot, R. (2017). *Accelerating the Future: The Economic Impact of the Emerging Passenger Economy.* Strategy Analytics Report. https://newsroom.intel.com/newsroom/wp-content/uploads/sites/11/2017/05/passenger-economy.pdf

Lane, C. (2013). Taste makers in the “fine-dining” restaurant industry: The attribution of aesthetic and economic value by gastronomic guides. *Poetics*, *41*(4), 342–365. doi:10.1016/j.poetic.2013.05.003

Lange, G. S., & Johnston, W. J. (2020). The value of business accelerators and incubators–an entrepreneur’s perspective. *Journal of Business and Industrial Marketing*.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Langleben, D. D., Loughead, J. W., Ruparel, K., Hakun, J. G., Bush-Winokur, S., Holloway, M. B., Strasser, A. A., Cappella, J. N., & Lerman, C. (2009). Reduced prefrontal and temporal processing and recall of high "sensation value" ads. *NeuroImage*, *46*(1), 219–225. doi:10.1016/j.neuroimage.2008.12.062 PMID:19457412

Langley. (n.d.). *Crafting Papers on Machine Learning.* Adaptive Systems Group, Daimler-Chrysler Research and Technology Center.

Langley, D. J., van Doorn, J., Ng, I. C., Stieglitz, S., Lazovik, A., & Boonstra, A. (2021). The Internet of Everything: *Smart things and their impact on business models. Journal of Business Research*, *122*, 853–863. doi:10.1016/j.jbusres.2019.12.035

Lara, F., & Deckers, J. (2019). Artificial Intelligence as a Socratic Assistant for Moral Enhancement. *Neuroethics*, 1–13. doi:10.100712152-019-09401-y

Larson, H., Jarrett, C., Eckersberger, E., Smith, D., & Paterson, P. (2014). Understanding vaccine hesitancy around vaccines from a global perspective: A systematic review of published literature, 2007-2012. *Vaccine*, *32*(19), 2150–2159. doi:10.1016/j.vaccine.2014.01.081 PMID:24598724

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, *6*(4), 239–242. doi:10.100712599-014-0334-4

Laudon, K. C., & Laudon, J. P. (2020). *Management information systems: managing the digital firm* (16th ed.). Pearson Education Limited.

Lauterbach, A. (2019). Artificial intelligence and policy: Quo vadis? *Digital Policy*. *Regulation & Governance*, *21*(3), 238–263. doi:10.1108/DPRG-09-2018-0054

Lazar, I., Motogna, S., & Parv, B. (2010). Behaviour-Driven Development of Foundational UML Components. Department of Computer Science. Babes-Bolyai University, Cluj-Napoca, Romania. doi:10.1016/j.entcs.2010.07.007

Leachman, S. A., & Merlino, G. (2017). Medicine: The Final Frontier in Cancer Diagnosis. *Nature*, *542*(7639), 36–38. doi:10.1038/nature21492 PMID:28150762

Lea, T.-H., Nguyen, C. P., & Park, D. (2020). Financing renewable energy development: Insights from 55 countries. *Energy Research & Social Science*, *68*, 101537. doi:10.1016/j.erss.2020.101537

Leavy, S. (2018, May). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In *Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14-16). 10.1145/3195570.3195580

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Lee, J., Singh, J., & Azamfar, M. (2019). *Industrial artificial intelligence.* arXiv preprint arXiv:1908.02150.

Lee, C.-L. (2014). Is co-branding a double-edged sword for brand partners? *European Research Studies Journal*, *17*(4), 19–34. doi:10.35808/ersj/430

Lee, H. L. (2018). Big data and the innovation cycle. *Production and Operations Management*, *27*(9), 1642–1646. doi:10.1111/poms.12845

Lee, I., & Shin, Y. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, *63*(2), 157–170. doi:10.1016/j.bushor.2019.10.005

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lee, M., Yun, J. J., Pyka, A., Won, D., Kodama, F., Schiuma, G., Park, H. S., Jeon, J., Park, K. B., Jung, K. H., Yan, M.-R., Lee, S. Y., & Zhao, X. (2018). How to Respond to the Fourth Industrial Revolution or the Second Information Technology Revolution? Dynamic New Combinations between Technology, Market, and Society through Open Innovation. *Journal of Open Innovation*, *4*(3), 21. Advance online publication. doi:10.3390/joitmc4030021

Lee, S. M., Tan, X., & Trimi, S. (2005). Current practices of leading e-government countries. *Communications of the ACM*, *48*(10), 99–104. doi:10.1145/1089107.1089112

Lee, T. (2019). Financial investment for the development of renewable energy capacity. *Energy & Environment*, *0*(0), 1–14. doi:10.1177/0958305X19882403

Legardinier, A. (2013). *Comment limiter les biais liés au choix des échelles de mesure dans les études marketing?* [How to limit the biases linked to the choice of measurement scales in marketing studies?]. Gestion et management, Dumas.

Legg, S., & Hutter, M. (2007). *Universal intelligence: a definition of machine intelligence*. Retrieved from https://arxiv.org/pdf/0712.3329.pdf

Leitch, R., & Day, Ch. (2000). *Action research and reflective practice: towards a holistic view*. Taylor & Francis. https://www.tandfonline.com/doi/ref/10.1080/09650790000200108?scroll=top

Leminen, S., Westerlund, M., & Nyström, A.-G. (2012). Living Labs as Open-Innovation Networks. *Technology Innovation Management Review*, *2*(9), 6–11. doi:10.22215/timreview/602

Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, *80*(6), 69–96. doi:10.1509/jm.15.0420

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*, 69–96.

Letheren, K., Russell-Bennett, R., & Whittaker, L. (2020). Black, white or grey magic? Our future with artificial intelligence. *Journal of Marketing Management*, *36*(3-4), 216–232. doi:10.1080/0267257X.2019.1706306

Leung, E., Paolacci, G., & Puntoni, S. (2018). Man versus machine: Resisting automation in identity-based consumer behavior. *JMR, Journal of Marketing Research*, *55*(6), 818–831. doi:10.1177/0022243718818423

Lever, M. W., Shen, Y., & Joppe, M. (2019). Reading travel guidebooks: Readership typologies using eye-tracking technology. *Journal of Destination Marketing & Management*, *14*, 1–13. doi:10.1016/j.jdmm.2019.100368

Levinson, M. (2020). *Outside The Box*. Kirkus Reviews.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal*, *9*, 181–212. doi:10.28945/479

Lewis, R. C. (1981). The positioning statement for hotels. *The Cornell Hotel and Restaurant Administration Quarterly*, *22*(1), 51–61. doi:10.1177/001088048102200111

Lewkowitz, M. (2014, Feb 12). *Bots: The future of human-computer interaction*. Available: https://chatbotsmagazine.com/bots-the-future-of-human-computer-interaction56696f7aff56

Liang, Z. (2019, September). Design of automatic matching system for ocean-going cargo in international logistics. *Journal of Coastal Research, 93*(SI), 1105–1110. doi:10.2112/SI93-160.1

Liban, A., & Hilles, S. M. (2018, July). Latent Fingerprint Enhancement Based on Directional Total Variation Model with Lost Minutiae Reconstruction. In *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)* (pp. 1-5). IEEE. 10.1109/ICSCEE.2018.8538417

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Li, H., Xie, K. L., & Zhang, Z. (2020). The effects of consumer experience and disconfirmation on the timing of online review: Field evidence from the restaurant business. *International Journal of Hospitality Management*, *84*, 102344. doi:10.1016/j.ijhm.2019.102344

Li, J. P., Mirza, N., Rahat, B., & Xiong, D. (2020). Machine learning and credit ratings prediction in the age of fourth industrial revolution. *Technological Forecasting and Social Change*, *161*, 120309. doi:10.1016/j.techfore.2020.120309

Li, J., Feng, J., & Kuo, C. C. J. (2018). Deep convolutional neural network for latent fingerprint enhancement. *Signal Processing Image Communication*, *60*, 52–63. doi:10.1016/j.image.2017.08.010

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323. doi:10.1016/j.tourman.2018.03.009

Likert, R. (1932). A technique for the measurement of attitudes. *Archives de Psychologie*.

Lim, C., Kim, K., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities (London, England)*, *82*, 86–99. doi:10.1016/j.cities.2018.04.011

Lin, A., & Chen, N. C. (2012). Cloud computing as an innovation: Percepetion, attitude, and adoption. *International Journal of Information Management*, *32*(6), 533–540. doi:10.1016/j.ijinfomgt.2012.04.001

Lin, C., & Lekhawipat, W. (2014). Factors affecting online repurchase intention. *Industrial Management & Data Systems*, *114*(4), 597–611. doi:10.1108/IMDS-10-2013-0432

Lind, M. (2018). Digital Data Sharing: The Ignored Opportunity for Making Global Maritime Transport Chains More Efficient | UNCTAD. *UNCTAD Transport and Trade Facilitation Newsletter, 79*.

Lind, M., Michaelides, M., Ward, R., & Watson, R. T. (2020). *Maritime Informatics*. Springer International Publishing. Retrieved from https://books.google.pt/books?id=Ed6PzQEACAAJ

Lind, M., Michaelides, M., Ward, R., & Watson, R. T. (2020). *Maritime Informatics*. Springer International Publishing.

Lindstrom, M. (2016). *Small data: las pequeñas pistas que nos advierten de las grandes tendencias*. Grupo Planeta Spain.

Lindstrom, M. (2019). *Brandwashed: O lado oculto do marketing controlamos o que compramos ou são as empresas que escolhem por nós?* Alta Books.

Linke, M. (2012). Connecting Logistics Networks Globally Via the Un Single Window Concept. *Journal of Globalization Studies*, *3*(2), 139–154.

Lin, P., Abney, K., & Bekey, G. (2011). Robot Ethics: Mapping the Issues for a Mechanized World. *Artificial Intelligence*, *175*(5-6), 942–949. doi:10.1016/j.artint.2010.11.026

Lin, P., Abney, K., & Bekey, G. A. (Eds.). (2012). *Robot ethics: The ethical and social implications of robotics*. MIT Press.

Lin, Y., Osman, M., & Ashcroft, R. (2017). Nudge: Concept, effectiveness, and ethics. *Basic and Applied Social Psychology*, *39*(6), 293–306. doi:10.1080/01973533.2017.1356304

Lippert, S. K., & Govindarajulu, C. (2006). Technological, organizational, and environmental antecedents to web services adoption. *Communications of the IIMA*, *6*(1), 14.

Liu, H., Jayawardhena, C., Dibb, S., & Ranaweera, C. (2019). Examining the trade-off between compensation and promptness in eWOM-triggered service recovery: A restorative justice perspective. *Tourism Management*, *75*, 381–392. doi:10.1016/j.tourman.2019.05.008

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, M., Liu, S., & Yan, W. (2018). Latent fingerprint segmentation based on ridge density and orientation consistency. *Security and Communication Networks*, *2018*, 2018. doi:10.1155/2018/4529652

Liuqu, Y., Fan, X., & Fu, P. L. (2015). From Customer Satisfaction to Customer Experience: Online Customer Satisfaction Practice in International E-commerce. In *International Conference on Cross-Cultural Design* (pp. 80-89). Springer. DOI: 10.1007/978-3-319-20934-0_8

Liu, S., Liu, M., & Yang, Z. (2017). Sparse coding-based orientation estimation for latent fingerprints. *Pattern Recognition*, *67*, 164–176. doi:10.1016/j.patcog.2017.02.012

Liu, W., & Chang, C. (2007). Variants of Principal Components Analysis. *2007 IEEE International Geoscience and Remote Sensing Symposium*, 1083-1086, 10.1109/IGARSS.2007.4422989

Li, X., Yen, C.-L., & Uysal, M. (2014). Differentiating with brand personality in economy hotel segment. *Journal of Vacation Marketing*, *20*(4), 323–333. doi:10.1177/1356766714527965

Lobera, J., Fernández Rodríguez, C. J., & Torres-Albero, C. (2020). Privacy, values and machines: Predicting opposition to artificial intelligence. *Communication Studies*, *71*(3), 448–465. doi:10.1080/10510974.2020.1736114

Lockwood, R. (2018). *Introduction The Relational Data Model*. Accessed and reviewed in April 2021, http://www.jakobsens.dk/Nekrologer.htm

Loginovskiy, O. V., Dranko, O. I., & Hollay, A. V. (2018). *Mathematical Models for Decision-Making on Strategic Management of Industrial Enterprise in Conditions of Instability*. In Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications (EMIT 2018), Leipzig, Germany.

London, A. J. (2019). Artificial intelligence and black-box medical decisions: Accuracy versus explainability. *The Hastings Center Report*, *49*(1), 15–21. doi:10.1002/hast.973 PMID:30790315

Long, G. (2007). *Transmedia Storytelling. Business, Aesthetics and Production at the Jim Henson Company* (Master Thesis). Massachusetts Institute of Technology.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *The Journal of Consumer Research*, *46*(4), 629–650. doi:10.1093/jcr/ucz013

Longoni, C., & Cian, L. (2020). Artificial intelligence in utilitarian vs. hedonic contexts: The "word-of-machine" effect. *Journal of Marketing*. doi:10.1177/0022242920957347

Lorenzo Romero, C., Alarcón de Amo, M. C., & Gómez Borja, M. A. (2011). Adopción de redes sociales virtuales: Ampliación del modelo de aceptación tecnológica integrando confianza y riesgo percibido. *Cuadernos de Economía y Dirección de la Empresa*, *14*, 194–205.

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2020). Artificial Intelligence in Business: State of the Art and Future Research Agenda. *Journal of Business Research*, *129*, 911–926. doi:10.1016/j.jbusres.2020.11.001

Loureiro, S. M. C., Japutra, A., Molinillo, S., & Bilro, R. G. (2021). Stand by me: Analyzing the tourist–intelligent voice assistant relationship quality. *International Journal of Contemporary Hospitality Management*. Advance online publication. doi:10.1108/IJCHM-09-2020-1032

Lourido Gómez, S., & Otero Neira, M.C. (2016). La dimensión afectiva de la experiencia de marca y el papel de las emociones en el comportamiento del consumidor. 26 *Congreso ACEDE Vigo, ponencia 3592ª*: [libro de actas].

Lun, Y. V., & Quaddus, M. A. (2009). An empirical model of the bulk shipping market. *International Journal of Shipping and Transport Logistics*, *1*(1), 37. Advance online publication. doi:10.1504/IJSTL.2009.021975

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, *38*(6), 937–947. doi:10.1287/mksc.2019.1192

Lwakatare, L. E., Raj, A., Bosch, J., Olsson, H. H., & Crnkovic, I. (2019). *A taxonomy of software engineering challenges for machine learning systems: An empirical investigation.* Paper presented at the International Conference on Agile Software Development. 10.1007/978-3-030-19034-7_14

Lwakatare, L. E., Raj, A., Crnkovic, I., Bosch, J., & Olsson, H. H. (2020). Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions. *Information and Software Technology*, *127*, 106368. doi:10.1016/j.infsof.2020.106368

Lyon, D. (2014). Surveillance, snowden, and big data: Capacities, consequences, critique. *Big Data & Society*, *1*(2), 1–13. doi:10.1177/2053951714541861

Maaß, D., Spruit, M., & de Waal, P. (2014). Improving short-term demand forecasting for short-lifecycle consumer products with data mining techniques. *Decision Analysis*, *1*(1), 4. doi:10.1186/2193-8636-1-4

Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(86), 2579–2605.

Máchová, R., & Lnenicka, M. (2015). Reframing e-government development indices with respect to new trends in ict. *Review of Economic Perspectives*, *15*(4), 383–412. doi:10.1515/revecp-2015-0027

MacInnis, D. J., & Folkes, V. S. (2010). The disciplinary status of consumer behavior: A sociology of science perspective on key controversies. *The Journal of Consumer Research*, *36*(6), 899–914. doi:10.1086/644610

Mackenzie, A. (2019). From API to AI: Platforms and their opacities. *Information Communication and Society*, *22*(13), 1989–2006. doi:10.1080/1369118X.2018.1476569

Maeda, H. (2015). Response option configuration of online administered Likert scales. *International Journal of Social Research Methodology*, *18*(1), 15–26. doi:10.1080/13645579.2014.885159

Magistretti, S., Dell'Era, C., & Petruzzelli, A. M. (2019). How intelligent is Watson? Enabling digital transformation through artificial intelligence. *Business Horizons*, *62*(6), 819–829. doi:10.1016/j.bushor.2019.08.004

Magnani, L., & Bardone, E. (2008). Distributed morality: Externalizing ethical knowledge in technological artifacts. *Foundations of Science*, *13*(1), 99–108. doi:10.100710699-007-9116-5

Maguire, M. (2001). Methods to support human-centred design. *Human-Computer Studies*, *55*, 587–634.

Mahanti, A., & Alhajj, R. (2005). Visual Interface for Online Watching of Frequent Itemset Generation in Apriori and Eclat. In *Proceedings of the Fourth International Conference on Machine Learning and Applications*. IEEE Computer Society. 10.1109/ICMLA.2005.68

Ma, J., Wang, A., Lin, F., Wesarg, S., & Erdt, M. (2019). A novel robust kernel principal component analysis for nonlinear statistical shape modeling from erroneous data. *Computerized Medical Imaging and Graphics*, *77*, 101638. doi:10.1016/j.compmedimag.2019.05.006 PMID:31550670

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, *120*, 262–273. doi:10.1016/j.jbusres.2020.07.045

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS One*, *13*(3), e0194889. Advance online publication. doi:10.1371/journal.pone.0194889 PMID:29584784

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing–Connecting computing power to human insights. *International Journal of Research in Marketing*, *37*(3), 481–504. doi:10.1016/j.ijresmar.2020.04.005

Malhotra, N. K. (2004). *Marketing Research an Applied Orientation* (Vol. 4). Pearson, Prentice Hall.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns: The construct, the scale, and a causal model. *Information Systems Research*, *15*(4), 336–355. doi:10.1287/isre.1040.0032

Malik, M. E., Naeem, B., & Munawar, M. (2012). Brand image: Past, present and future. *Journal of Basic and Applied Scientific Research*, *2*(12), 13069–13075.

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015), Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *HRI'15: Proceedings of the Tenth Annual ACM/ IEEE International Conference on Human-Robot Interaction*. New York, NY: ACM. 10.1145/2696454.2696458

Malle, B. F. (2016). Integrating Robot Ethics and Machine Morality: The Study and Design of Moral Competence in Robots. *Ethics and Information Technology*, *18*(4), 243–256. doi:10.100710676-015-9367-8

Malone, A., Caceres Cabana, Y. D., & Taya Zegarra, A. (2020). Informal food systems and differential mobility during the COVID-19 pandemic in Arequipa, Perú. *The Town Planning Review*. Advance online publication. doi:10.3828/tpr.2020.61

Manohlia, D. (2015). Review: 'Ex Machina', a Mogul Fashions the Droid of his Dreams. *The New York Times*.

Manovich, L. (2005). *Soft Cinema. The Language of New Media*. M.I.T. Press.

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., & Sanghvi, S. (2017). *What the future of work will mean for jobs, skills, and wages.* McKinsey Global Institute Report. https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages

Marbán, O., Segovia, J., Menasalvas, E., & Fernández-Baizán, C. (2009). Toward data mining engineering: A software engineering approach. *Information Systems*, *34*(1), 87–107. doi:10.1016/j.is.2008.04.003

Marcus, G. E., & Fisher, M. M. (1999). *Anthropology as Cultural Critique: An experimental moment in the human sciences.* doi:10.7208/chicago/9780226229539.001.0001

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Vintage.

Maridueña, M. R., Leyva, M., & Febles, A. (2016). Modelado y análisis de indicadores de ciencia y tecnología mediante mapas cognitivos difusos. *Ciencias de la Información*, *47*(1), 17–24.

Marketing Directo. (2019). *Las 4 tendencias en marketing digital que definirán el año 2020*. Retrieved December 18, 2020, from https://www.marketingdirecto.com/marketing-general/tendencias/las-4-tendencias-en-marketing-digital-que-definiran-el-ano-2020

Marr, B. (2017). The Complete Beginners' Guide to Artificial Intelligence. *Forbes*.

Martelo, R. J., Jimenez-Pitre, I., & Moncaris Gonzalez, L. (2017). Guía Metodológica para el Mejoramiento del Desarrollo de Software a través de la Aplicación de la Técnica Árboles de Problemas. *Información Tecnológica*, *28*(3), 87–94. doi:10.4067/S0718-07642017000300010

Martimo, P. (2017). *Disruptive Innovation and Maritime Sector - Discovering smart-shipping's potential to disrupt shipping*. Report Turku School of Economics.

Martín Lineros, E. (2020). *Webinar Alumni: La revolución del 5G*. Retrieved December 18, 2020, from https://www.youtube.com/watch?v=en8SzXmYmZU&feature=youtu.be

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Martinez Coral, P. (2017). "Seguro mató a confianza": Challenges for the adoption of digital government in Colombia. *Academic Journals: Inclusion and Development*, *5*(1), 63–72. doi:10.26620/uniminuto.inclusion.5.1.2018.63-72

Martinez Dominguez, M. (2018). Access to and use of information and communication technologies in Mexico: determinants. *PAAKAT: Journal of Technology and Society, 8*(14). Retrieved from . doi:10.32870/Pk.a8n14.316

Martínez Pérez, S., Fernández Robles, B., & Barroso Osuna, J. (2021). Augmented reality as a resource for training in higher education [La realidad aumentada como recurso para la formación en la educación superior]. *Campus Virtuales*, *10*(1), 9–19.

Martínez Sánchez, F. (2008). La emoción. In Motivación & emoción. Mc Graw Hill.

Martínez-López, F. J., García-Ordaz, M., Arteaga-Sánchez, R., & Infante-Moro, A. (2015). The presence of large Spanish companies in online social networks. *Journal of Marketing Analytics*, *3*(4), 171–186. doi:10.1057/jma.2015.15

Martinho-Truswell, E. (2018), 3 Questions About AI That Nontechnical Employees Should Be Able to Answer. *Harvard Business Review*. https://hbr.org/2018/08/3-questionsabout-ai-that-nontechnicalemployees-should-be-able-to-answer

Mascarenhas, O. A., Kesavan, R., & Bernacchi, M. (2006). Lasting customer loyalty: A total customer experience approach. *Journal of Consumer Marketing*.

Masicampo, E. J., & Baumeister, R. F. (2011). Consider it done! Plan making can eliminate the cognitive effects of unfulfilled goals. *Revista de personalidad y psicología social, 101*(4), 667-683. doi:10.1037/a0024192

Mateas, M. (2003). *Expressive AI Games and AI, In proceedings of Level UP.* Digital Games Research Conference, Utrecht.

Mateas, M., & Murray, J. (2004). A preliminary poetics for interactive drama & games, & from game story to cyberdrama. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First Person: New Media as Story, Performance & Game* (pp. 2–33). The MIT Press.

Mathews, B. P., & Shepherd, J. L. (2002). Dimensionality of Cook and Wall's (1980) British Organizational Commitment Scale revisited. *Journal of Occupational and Organizational Psychology*, *75*(3), 369–375. doi:10.1348/096317902320369767

Matthew, C. T., & Sternberg, R. J. (2009). *Developing experience-based (tacit) knowledge through reflection. In Learning and Individual Differences.* Elsevier Science., doi:10.1016/j.lindif.2009.07.001

Matwin, S. (2013). Privacy-preserving data mining techniques: survey and challenges. In *Discrimination and Privacy in the Information Society* (pp. 209–221). Springer. doi:10.1007/978-3-642-30487-3_11

Maughan, L., Gutnikov, S., & Stevens, R. (2007). Like more, look more. Look more, like more: The evidence from eye tracking. *Journal of Brand Management*, *14*(4), 335–342. doi:10.1057/palgrave.bm.2550074

Mavragani, A., & Tsagarakis, K. P. (2019). Predicting referendum results in the Big Data Era. *Journal of Big Data*, *6*(1), 1–20. doi:10.118640537-018-0166-z

Mazuryk, T., & Gervautz, M. (1999). *History, Applications, Technology and Future*. Institute of Computer Graphics, Vienna University of Technology.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68. PMID:23074865

McAteer, M. (2020). *Nitpicking machine learning technical debt*. Retrieved from https://matthewmcateer.me/blog/machine-learning-technical-debt/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955), A proposal for the Dartmouth summer research project on artificial intelligence, http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

McCarthy, J. (1960). Recursive Functions of Symbolic Expressions and their Computation by Machine, part I. *Communications of the ACM*, *3*(4), 184–195. doi:10.1145/367177.367199

McCarthy, J. (1990). *What is AI? Basic Questions; Artificial Intelligence, Logic and Formalizing Common Sense*. Computer Science Department, Stanford University.

McCauley, L. (2007). *Countering the Frankenstein Complex.* American Association for Artificial Intelligence. https://www.openculture.com/2019/03/isaac-asimov-predicts-the-future-of-civilization.html

McCauley, L. (2007). *The Frankenstein Complex and Asimov's Three Laws*. Association of the Advancement of Artificial Intelligence.

McKinsey. (2020). *From surviving to thriving: Business after coronavirus*. Author.

McQuillan, D. (2018). People's councils for ethical machine learning. *Social Media + Society*, *4*(2), 2056305118768303. doi:10.1177/2056305118768303

Medallia & DEC. (2020*). ¿Pueden ser emocionales las empresas financieras?* Retrieved December 18, 2020, from https://go.medallia.com/es-dec-webinar.html

Mediamarketing. (2017). *Le chemin passe par le neuromarketing* [The path goes through neuromarketing]. Accessed from: https://www.mediamarketing.ma/article/GHBGBGH/le_chemin_passe_par_le_neuromarketing.html

Mehari, Y. (2020). *The Role of Social Trust in Citizen Mobility During COVID-19*. https://ssrn.com/abstract=3607668[REMOVED HYPERLINK FIELD] doi:10.2139srn.3607668

Mehra, A., Grundy, J., & Hosking, J. (2005). A generic approach to supporting diagram differencing and merging for collaborative design. In *ASE '05 Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering*. ACM. 10.1145/1101908.1101940

Mehrara, M., Rezaei, S., & Razi, H. D. (2015). Determinants of renewable energy consumption among ECO countries; based on Bayesian model averaging and weighted-average least square. *Int Lett Soc Hum Sci*, *54*, 96–109. doi:10.18052/www.scipress.com/ILSHS.54.96

Mehrtens, J., Cragg, P. B., & Mills, A. M. (2001). A model of Internet adoption by SMEs.[. *Information & Management*, *39*(3), 165–176. doi:10.1016/S0378-7206(01)00086-6

Mejia, J., Mankad, S., & Gopal, A. (2020). Service Quality Using Text Mining: Measurement and Consequences. *Manufacturing & Service Operations Management*, 1–19. doi:10.1287/msom.2020.0883

Meltzer, T. (2014, June 15). Robot doctors, online lawyers and automated architects: The future of the professions. *The Guardian*. https://www.theguardian.com/technology/2014/jun/15/robot-doctors-

Mendes Constante, J., Langen, P., Vieira, G., Lunkes, R., & van der Lugt, L. (2018). The impact of management practices use on Brazilian Port Authorities' Performance. *Rivista Internazionale di Economia dei Trasporti*, *45*. Advance online publication. doi:10.19272/201806702005

Méndez Aparicio, M. D. (2019). *Determinantes y consecuencias de la Experiencia de Cliente y la Satisfacción en el ámbito digital: un análisis del uso de áreas privadas*© (Doctoral thesis). Universidad de Burgos, Burgos.

Méndez-Aparicio, M.D., Jiménez-Zarco, A.I., Izquierdo-Yusta, A., & Martinez-Ruiz. (2020). The psychological logic of customer loyalty: towards an integrated conceptual framework (satisfaction, trust and commitment). *Publicación en curso.*

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Méndez-Aparicio, M. D., Jiménez-Zarco, A., Izquierdo-Yusta, A., & Blazquez-Resino, J. J. (2020). Customer Experience and Satisfaction in Private Insurance Customer Web Areas. *Frontiers in Psychology*, *11*, 2591.

Mengesha, N., Ayanso, A., & Demissie, D. (2020). Profiles and Evolution of E-Government Readiness in Africa: A Segmentation Analysis. *International Journal of Information Systems and Social Change*, *11*(1), 43–65. doi:10.4018/IJISSC.2020010104

Menon, R. R., Sukhadiya, H., & Patel, J. (2015). Standalone USB Flash to USB Flash. *Data Transfer*, ●●●, 6936–6941.

Menz, F. C., & Vachon, S. (2006). The effectiveness of different policy regimes for promoting wind power: Experiences from the states. *Energy Policy*, *34*(14), 1786–1796. doi:10.1016/j.enpol.2004.12.018

Menzies, T. (2019). The five laws of SE for AI. *IEEE Software*, *37*(1), 81–85. doi:10.1109/MS.2019.2954841

Merisavo, M., Kajalo, S., Karjaluoto, H., Virtanen, V., Salmenkivi, S., Raulas, M., & Leppäniemi, M. (2007). An empirical study of the drivers of consumer acceptance of mobile advertising. *Journal of Interactive Advertising*, *7*(2), 1–19. doi:10.1080/15252019.2007.10722130

Merritt, S. M. (2012). The two-factor solution to Allen and Meyer's (1990) Affective Commitment Scale: Effects of negatively worded items. *Journal of Business and Psychology*, *27*(4), 421–436. doi:10.100710869-011-9252-3

Merson, P. (2009). *Data Model as an Architectural View. Technical Note. CMU/SEI-2009-TN-024*. Research, Technology, and System Solutions.

Microsoft. (2020a). *What are Azure Machine Learning pipelines?* Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/concept-ml-pipelines

Microsoft. (2020b). *What is the Team Data Science Process?* Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/overview

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261–276. doi:10.1016/j.jbusres.2019.01.044

Mileski, J., Clott, C., Galvao, C. B., & Laverne, T. (2020). Technical analysis: the psychology of the market of dry bulk freight rates. *Journal of Shipping & Trade, 5*(1). doi:10.118641072-020-00079-7

Miller, K. (2020). A Matter of Perspective: Discrimination, Bias, and Inequality in AI. In M. Jackson & M. Shelly (Eds.), Legal Regulations, Implications, and Issues Surrounding Digital Data (pp. 182-202). IGI Global.

Miller, G. (2018). *Why Altered Carbon is not about the future-nor is any other science fiction*. The Conversation.

Millington, I. (2006). Artificial Intelligence for Games. Elsevier, Morgan Kaufmann Publishers.

Ministério da Infraestrutura. (2018). Porto sem Papel - PSP — Português (Brasil). Author.

Ministério da Infraestrutura. (2020). Sistema Portuário Nacional — Português (Brasil). *Governo Do Brasil*. Retrieved from https://www.gov.br/infraestrutura/pt-br/assuntos/transporte-aquaviario/sistema-portuario

Mintz, Y., & Brodie, R. (2019). Introduction to artificial intelligence in medicine. *Minimally Invasive Therapy & Allied Technologies*, *28*(2), 73–81. doi:10.1080/13645706.2019.1575882 PMID:30810430

Miranda, J., Lopez, C. S., Navarro, S., Bustamante, M. R., Molina, J. M., & Molina, A. (2019). Open Innovation Laboratories as Enabling Resources to Reach the Vision of Education 4.0. *2019 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC), Engineering, Technology and Innovation (ICE/ITMC), 2019 IEEE International Conference On*. 10.1109/ICE.2019.8792595

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mircioiu, C., & Atkinson, J. (2017). A comparison of parametric and non-parametric methods applied to a Likert scale. *Pharmacy (Basel, Switzerland)*, *5*(2), 26. doi:10.3390/pharmacy5020026 PMID:28970438

Missura, O., & Gartner, T. (2009). *Player modelling for Intelligent Difficulty Adjustment*. Discovery Science, 12th International Conference.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). *Model cards for model reporting.* Paper presented at the Conference on Fairness, Accountability, and Transparency. 10.1145/3287560.3287596

Modelling, T. O. G. A. F. (2015a). *Data dissemination view*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-dissemination-diagrams-menu.html

Modelling, T. O. G. A. F. (2015c). *Data lifecyle diagram*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-lifecycle-diagrams-menu.html

Modelling, T. O. G. A. F. (2015d). *Data migration diagram*. TOGAF Modelling. Accessed and reviewed in April 2021, http://www.TOGAF-modelling.org/models/data-architecture-menu/data-migration-diagrams-menu.html

Modelling, T. O. G. A. F. (2015e). *Data security diagram*. TOGAF Modelling. Accessed and reviewed in April 2021,http://www.TOGAF-modelling.org/models/data-architecture-menu/19-data-security-diagrams.html

Moeuf, A., Lamouri, S., Pellerin, R., Tamayo-Giraldo, S., Tobon-Valencia, E., & Eburdy, R. (2020). Identification of critical success factors, risks and opportunities of Industry 4.0 in SMEs. *International Journal of Production Research*, *58*(5), 1384–1400. doi:10.1080/00207543.2019.1636323

Moher, D., Liberati, A., & Tetzlaff, J. (2015). Principais itens para relatar revisões sistemáticas e meta-análises: A recomendação PRISMA. *Epidemiologia e Serviços de Saúde: Revista do Sistema Unico de Saúde do Brasil*, *24*(2), 335–342. doi:10.5123/S1679-49742015000200017

Molnar, C. (2020). *Interpretable machine learning*. Lulu.Com.

Moon, H., & Sprott, D. E. (2016). Ingredient branding for a luxury brand: The role of brand and product fit. *Journal of Business Research*, *69*(12), 5768–5774. doi:10.1016/j.jbusres.2016.04.173

Moon, K. S., & Kim, H. (2019). Performance of deep learning in prediction of stock market volatility. *Economic Computation and Economic Cybernetics Studies and Research*, *53*(2/2019), 77–92. doi:10.24818/18423264/53.2.19.05

Moore, J. (2014). *Java programming with lambda expressions-A mathematical example demonstrates the power of lambdas in Java 8*. Javaworld. Accessed and reviewed in April 2021, https://www.javaworld.com/article/2092260/java-se/java-programming-with-lambda-expressions.html

Mora-Fernandez, J. I. (2019a). Extended analysis of the transmedia narrative universe of Nolan's Dark Knight: integrating original authorships, media convergences & interactive linear narratives. In Reflections in Social Sciences and other Topics. Guayaquil: CIDE.

Mora-Fernandez, J. I. (2019b). Concepts and Models of Analysis of Interactive and Transmedia Narratives: A Batman's Universe Case Study. In *Information Technology and Systems ICITS 2019, AISC 918* (pp. 929–943). Springer. doi:10.1007/978-3-030-11890-7_87

Mora-Fernandez, J. I. (2020). IMotions' Automatic Facial Recognition & Text-Based Content Analysis of Basic Emotions & Empathy in the Application of the Interactive Neurocommunicative Technique LNCBT (Line & Numbered Concordant Basic Text). In Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Communication, Organization and Work. Springer. https://doi.org/10.1007/978-3-030-49907-5_5.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mora-Fernández, J.I. (2009). *La interfaz hipermedia: el paradigma de la comunicación interactiva. Modelos para implementar la inmersión juvenil en multimedia interactivos culturales. (Videojuegos, cine, realidad aumentada, museos y web).* SGAE. Fundación Autor. Ediciones Autor.

Mora-Fernandez, J. I. (2013). Artecnología en cine interactivo al unas categorías, interfaces, estructuras narrativas, emociones e investigaciones. In *ArTecnologia. Arte, Tecnologia e linguagens Midiáticas*. Buqui.

Morales Salas, R. E., Infante-Moro, J. C., & Gallardo-Pérez, J. (2019). Mediation and interaction in a VLE for effective management in virtual learning [La mediación e interacción en un AVA para la gestión eficaz en el aprendizaje virtual]. *Campus Virtuales*, *8*(1), 49–61.

Morales Salas, R. E., Infante-Moro, J. C., & Gallardo-Pérez, J. (2020). Evaluation of virtual learning environments: A management to improve. *IJERI: International Journal of Educational Research and Innovation*, (13), 126–142.

Morales Salas, R. E., & Rodríguez Pavón, P. R. (2020). Digital ICT skills applied in organizations [Las competencias digitales en TIC aplicadas en las organizaciones]. *International Journal of Information Systems and Software Engineering for Big Companies*, *7*(1), 25–35.

Moreno Sánchez, I. (2002). *Musas y Nuevas Tecnologías: El relato hipermedia*. Paidós Comunicación.

Morgan, B. (2018). How Amazon has reorganized around artificial intelligence and machine learning. *Forbes*. https://www.forbes.com/sites/blakemorgan/2018/07/16/how-amazon-has-re-organized-around-artificial-intelligence-and-machine-learning/?sh=40ca07877361

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, *58*(Julio), 20–38.

Morris, P. L., Hopwood, M., Whelan, G., Gardiner, J., & Drummond, E. (2001). Naltrexone for alcohol dependence: A randomized controlled trial. *Addiction (Abingdon, England)*, *11*(15), 65–73. doi:10.1046/j.1360-0443.2001.961115654.x PMID:11784454

Mostert, P., Petzer, D., & De Meyer, C. (2012). A theoretical and empirical investigation into service failure and service recovery in the restaurant Industry. In Service Science Research, Strategy and Innovation: Dynamic Knowledge Management Methods (pp. 86-99). IGI Global. doi:10.4018/978-1-4666-0077-5.ch005

Motion, J., Leitch, S., & Brodie, R. J. (2003). Equity in corporate co-branding: The case of Adidas and the All Blacks. *European Journal of Marketing*, *37*(7/8), 1080–1094. doi:10.1108/03090560310477672

Mouratiadou, I., & Moran, D. (2007). Mapping public participation in the Water Framework Directive: A case study of the Pinios River Basin, Greece. *Ecological Economics*, *62*(1), 66–76. doi:10.1016/j.ecolecon.2007.01.009

Mukhopadhyay, P., & Chaudhuri, B. B. (2015). A survey of Hough Transform. *Pattern Recognition*, *48*(3), 993–1010. doi:10.1016/j.patcog.2014.08.027

Mullainathan, S., & Shafir, E. (2016). *Scarcity: Why does having little mean so much?* Fondo de Cultura Económica.

Müller, V. C., & Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In Fundamental Issues of Artificial Intelligence (555-572). Springer. doi:10.1007/978-3-319-26485-1_33

Munin. (2016). http://www.unmanned-ship.org/munin/

Murphy, R. R., & Woods, D. D. (2009). *Beyond Asimov: The Three Laws of Responsible Robotics, Institute for Human and Machine Cognition*. IEEE Computer Society.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*(5), 723–739. doi:10.1037/0022-3514.64.5.723 PMID:8505704

Murray, P. N. (2013). *Inside the Consumer Mind How Emotions Influence What We Buy The emotional core of consumer decision making*. Retrieved from https://www.psychologytoday.com/blog/inside-the-consumermind/201302/how-emotions-influence-what-we-buy

Murray, J. (1997). *Hamlet On the Holodeck: the Future of Narrative in Cyberspace*. Free Press.

Murray, J. (2012). *Inventing The Medium. Principles of Interaction, Design as a Cultural Practice*. The MIT Press.

Murray, J. (2013). Likert data: What to use, parametric or non-parametric? *International Journal of Business and Social Science*, *4*(11).

Mustak, M., Jaakkola, E., Halinen, A., & Kaartemo, V. (2016). Customer participation management. *Journal of Service Management*, *27*(3), 250–275. doi:10.1108/JOSM-01-2015-0014

Nack, F., & Gordon, A. S. (2016). Interactive Storytelling. ICIDS 2016, LNCS 10045, 61–72. DOI: 6 doi:10.1007/978-3-319-48279-8

Nantes, A., Brown, R., & Maire, F. (2008). A Framework for Semi- Automatic Testing of Video Games. In *Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference*. Association for the Advancement of Artificial Intelligence.

Nave, G., Nadler, A., Dubois, D., Zava, D., Camerer, C., & Plassmann, H. (2018). Single-dose testosterone administration increases men's preference for status goods. *Nature Communications*, *9*(1), 24–33. doi:10.103841467-018-04923-0 PMID:29970895

Nayar, R. (2015). Role of Web 3.0 in Service Innovation. In *The Handbook of Service Innovation* (pp. 253–280). Springer. doi:10.1007/978-1-4471-6590-3_13

Ndou, V., Secundo, G., Dumay, J., & Gjevori, E. (2018). Understanding intellectual capital disclosure in online media big data: An exploratory case study in a university. *Meditari Accountancy Research*, *26*(3), 499–530. doi:10.1108/MEDAR-03-2018-0302

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance through Text Mining. *Marketing Science*, *31*(3), 521–543. doi:10.1287/mksc.1120.0713

Neves, A. R., Costa, J., & Reis, J. (2021). Using a Systematic Literature Review to Build a Framework for University-Industry Linkages using Open Innovation. *Procedia Computer Science*, *181*, 23–33. doi:10.1016/j.procs.2021.01.095

Nguyen, A., & Catalan, D. (2020). Digital mis/disinformation and public engagement with health and science controversies: Fresh perspectives from Covid-19. *Media and Communication*, *8*(2), 323–328. doi:10.17645/mac.v8i2.3352

Nicholls, M. E., Orr, C. A., Okubo, M., & Loftus, A. (2006). Satisfaction guaranteed: The effect of spatial biases on responses to Likert scales. *Psychological Science*, *17*(12), 1027–1028. doi:10.1111/j.1467-9280.2006.01822.x PMID:17201782

Nielsen. (2019). *3,5 Mil Milhões de Euros na "Selva Promocional" dos Bens de Grande Consumo*. Nielsen. Retrieved from https://www.nielsen.com/pt/pt/insights/article/2019/3-point-5-billion-euros-in-promotional-jungle-of-consumer-goods/

Nielsen. (2020). *COVID-19 will affect low-income consumers in Latin America the most.* Retrieved from https://www.nielsen.com/pe/es/insights/article/2020/covid-19-afectara-mas-a-los-consumidores-de-bajos-ingresos-en-latinoamerica/

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nieto, A. B., Galindo-Villardón, M., Leiva, V., & Vicente-Galindo, M. (2014). A Methodology for Biplots based on bootstrapping with R. *Revista Colombiana de Estadistica*, *37*(2Spe), 367–397. doi:10.15446/rce.v37n2spe.47944

Nijboer, F., Morin, F., Carmien, S., Koene, R., Leon, E., & Hoffman, U. (2009). Affective brain-computer interfaces: Psychophysiological markers of emotion in healthy persons and in persons with amyotrophic lateral sclerosis. In *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE. 10.1109/ACII.2009.5349479

Nilashi, M., Samad, S., Ahmadi, N., Ahani, A., Abumalloh, R. A., Asadi, S., & Yadegaridehkordi, E. (2020). Neuromarketing: A review of research and implications for marketing. *Journal of Soft Computer Decision Support System*, *7*(2), 23–31.

Ning, L. Q., & Liu, J. L. (2008). Application of Stone-Geisser Test in Customer Satisfaction. *Tianjin Daxue Xuebao*, 3.

Nobre, H. (2019). A marca como ferramenta de gestão empresarial. In C. Machado & J. P. Davim (Eds.), *Organização e Políticas Empresariais* (pp. 133–166). Actual Editora.

Nobre, H., & Simões, C. (2019). NewLux brand relationship scale: Capturing the scope of mass-consumed luxury brand relationships. *Journal of Business Research*, *102*, 328–338. doi:10.1016/j.jbusres.2019.01.047

Nofz, M.P., & Vendy, P. (2002). When Computers Say It with Feeling: Communication and Synthetic Emotions in Kubrick's 2001: A Space Odyssey. *Journal of Communication Inquiry*, 25-45.

Nori, H., Jenkins, S., Koch, P., & Caruana, R. (2019). *Interpretml: A unified framework for machine learning interpretability.* arXiv preprint arXiv:1909.09223.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education: Theory and Practice*, *15*(5), 625–632. doi:10.100710459-010-9222-y PMID:20146096

Norstad, I., Fagerholt, K., & Laporte, G. (2011, August). Tramp ship routing and scheduling with speed optimization. *Transportation Research Part C, Emerging Technologies*, *19*(5), 853–865. doi:10.1016/j.trc.2010.05.001

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152–159. doi:10.1016/j.tics.2011.01.005 PMID:21376657

Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S. S., Reuter, U., Gama, J., & Gandomi, A. H. (2020b). Data Science in Economics: Comprehensive Review of Advanced Machine Learning and Deep Learning Methods. *Mathematics*, *8*(10), 1799. doi:10.3390/math8101799

Nosratabadi, S., Mosavi, A., Keivani, R., Ardabili, S., & Aram, F. (2020a). State of the Art Survey of Deep Learning and Machine Learning Models for Smart Cities and Urban Sustainability. In A. Várkonyi-Kóczy (Ed.), *Engineering for Sustainable Future. INTER-ACADEMIA 2019. Lecture Notes in Networks and Systems* (Vol. 101). Springer. doi:10.1007/978-3-030-36841-8_22

Notteboom, T., Lugt, L. van der, Saase, N. van, Sel, S., & Neyens, K. (2020). The Role of Seaports in Green Supply Chain Management: Initiatives, Attitudes, and Perspectives in Rotterdam, Antwerp, North Sea Port, and Zeebrugge. *Sustainability, 12*(4), 1688. doi:10.3390u12041688

Notteboom, T. E., Parola, F., Satta, G., & Pallis, A. A. (2017). The relationship between port choice and terminal involvement of alliance members in container shipping. *Journal of Transport Geography*, *64*, 158–173. doi:10.1016/j.jtrangeo.2017.09.002

Ntwali, A., Kituyi, A., & Kengere, A. O. (2020). Claims Management and Financial Performance of Insurance Companies in Rwanda: A Case of SONARWA General Insurance Company Ltd. *Journal of Financial Risk Management*, *9*(03), 190.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nueno, J. L., & Quelch, J. A. (1998). The mass marketing of luxury. *Business Horizons*, *41*(6), 61–61. doi:10.1016/S0007-6813(98)90023-4

Nunes, O., & Nisi, V. (2017). Interactive Storytelling. *ICIDS 2017, LNCS 10690*, 372–375. doi:10.1007/978-3-319-71027-3_46

Nunnally, J. C. (1978). *Psychometric Theory*. Academic Press.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.

O'Neil, J. M., Newton, R. J., Bone, E. K., Birney, L. B., Green, A. E., Merrick, B., Goodwin-Segal, T., Moore, G., & Fraioli, A. (2020). Using urban harbors for experiential, environmental literacy: Case studies of New York and the Chesapeake Bay. *Regional Studies in Marine Science*, *33*, 100886. Advance online publication. doi:10.1016/j.rsma.2019.100886

O'Reilly, L. (2013). *Google patents 'pay-per-gaze' eye-tracking ad technology*. Accessed from: https://www.marketingweek.com/google-patents-pay-per-gaze-eye-tracking-ad-technology/

OASIS. (2009). *OASIS Web Services Reliable Messaging (WS-ReliableMessaging) Version 1.2*. OASIS Standard. Accessed and reviewed in April 2021, http://docs.oasis-open.org/ws-rx/wsrm/200702/wsrm-1.2-spec-os.html

OASIS. (2014). *ISO/IEC and OASIS Collaborate on E-Business Standards-Standards Groups Increase Cross-Participation to Enhance Interoperability*. The OASIS Group. Accessed and reviewed in April 2021, https://www.oasis-open.org/news/pr/isoiec-and-oasis-collaborate-on-e-business-standards

Öberg, C. (2016). What creates a collaboration-level identity? *Journal of Business Research*, *69*(9), 3220–3230. doi:10.1016/j.jbusres.2016.02.027

Obschonka, M., & Audretsch, D. B. (2020). Artificial intelligence and big data in entrepreneurship: A new era has begun. *Small Business Economics*, *55*(3), 529–539. doi:10.100711187-019-00202-4

OECD. (2005). *E-government for better government*. OECD.

OECD. (2019). *OECD Public Governance Reviews The Innovation System of the Public Service of Brazil An Exploration of its Past, Present and Future Journey*. OECD Publishing.

Ogbum, W. F., & Thomas, D. S. (1922, May 8). The influence of the business cycle on certain social conditions. *Journal of the American Statistical Association*, *18*(139), 324–340. doi:10.1080/01621459.1922.10502475

Oh, H., Fiore, A. M., & Jeong, M. (2007). Measuring experience economy concepts: Tourism applications. *Journal of Travel Research*, *46*(2), 119–131.

Ohme, R. (2001). The implicit conditioning of consumer attitudes: Logo substitution effect. Implicit Processes in Cognition, 32(1).

Ohme, R., Matukin, M., & Wicher, P. (2020). Merging Explicit Declarations With Implicit Response Time to Better Predict Behavior. In V. Chkoniya, A. O. Madsen, & P. Bukhrashvili (Eds.), *Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior* (pp. 427–448). IGI Global. doi:10.4018/978-1-7998-3115-0.ch023

Okazaki, S. (2006). What do we know about mobile internet adopters? A cluster analysis. *Information & Management*, *43*(2), 127–141. doi:10.1016/j.im.2005.05.001

Oke, A., & Fernandes, F. A. P. (2020). Innovations in Teaching and Learning: Exploring the Perceptions of the Education Sector on the 4th Industrial Revolution (4IR). *Journal of Open Innovation*, *6*(2), 31. Advance online publication. doi:10.3390/joitmc6020031

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Okonkwo, U. (2009). The luxury brand strategy challenge. *Journal of Brand Management*, *16*(5-6), 287–289. doi:10.1057/bm.2008.53

Olajide, V. (2019). *Introduction to Data Management*. Accessed from: https://www.researchgate.net/publication/335062781

Oliva, F. L., & Kotabe, M. (2019). Barriers, practices, methods and knowledge management tools in startups. *Journal of Knowledge Management*.

Oliveira, B., & Casais, B. (2019). The importance of user-generated photos in restaurant selection. *Journal of Hospitality and Tourism Technology*, *10*(1), 2–14. doi:10.1108/JHTT-11-2017-0130

Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *JMR, Journal of Marketing Research*, *17*(4), 460–469. Advance online publication. doi:10.1177/002224378001700405

Olson, Bartley, Urbanowicz, & Moore. (2016). Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, *GECCO '16* (pp. 485-492). Association for Computing Machinery.

OMNI-SCI. (2021). *Data Science - A Complete Introduction*. OMNI-SCI. Accessed and reviewed in April 2021, https://www.omnisci.com/learn/data-science

Omri, A., Daly, S., & Nguyen, D. K. (2015). A robust analysis of the relationship between renewable energy consumption and its main drivers. *Applied Economics*, *47*(28), 2913–2923. doi:10.1080/00036846.2015.1011312

Oosterbaan, R. J. (2019). Software for generalized and composite probability distributions. *International Journal of Mathematical and Computational Methods*, *4*, 1–9.

Ordanini, A., & Pasini, P. (2008). Service co-production and value co-creation: The case for a service-oriented architecture (SOA). *European Management Journal*, *26*(5), 289–297. doi:10.1016/j.emj.2008.04.005

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206. doi:10.1016/j.actpsy.2013.06.003 PMID:23845447

Ortega-Vivanco, M. (2020, March). Effects of Covid-19 on consumer behaviour: The case of Ecuador. *RETOS. Journal of Management Science and Economics*, *10*(20). Advance online publication. doi:10.17163/ret.n20.2020.03

Osobajo, O. A., Koliousis, I., & McLaughlin, H. (2021). Making sense of maritime supply chain: A relationship marketing approach. *Journal of Shipping and Trade*, *6*(1), 1. doi:10.118641072-020-00081-z

Oullier, O. (2012). Le cerveau et la loi: analyse de l'émergence du neurodroit [The Brain and the Law: Analysis of the Emergence of Neurolaw]. Centre d'Analyse Stratégique.

Oved, D. (2018). *Real-time human pose estimation in the browser with TensorFlow.js*. Retrieved from https://medium.com/tensorflow/real-time-human-pose-estimation-in-the-browser-with-tensorflow-js-7dd0bc881cd5

Overbye, D. (2018). 2001: A Space Odyssey' Is Still the 'Ultimate Trip'. *The New York Times*. https://www.nytimes.com/2018/05/10/science/2001-a-space-odyssey-kubrick.html

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. *California Management Review*, *61*(4), 156–185. doi:10.1177/0008125619859318

Özesmi, U., & Özesmi, S. L. (2003). A participatory approach to ecosystem conservation: fuzzy cognitive maps and stakeholder group analysis in Uluabat Lake, Turkey. *Environmental Management*, *31*(4), 518-531. doi:10.100700267-002-2841-1

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Özesmi, U., & Özesmi, S. L. (2004). Ecological models based on people's knowledge: A multistep Fuzzy Cognitive Mapping approach. *Ecological Modelling*, *176*(1), 43–64. doi:10.1016/j.ecolmodel.2003.10.027

Ozkaya, I. (2020). What is really different in engineering AI-enabled systems? *IEEE Software*, *37*(4), 3–6. doi:10.1109/MS.2020.2993662

Öztürk, L. (2012). Public Investments and Regional Inequality: A Causality Analysis, 1975-2001. *Ege Academıc Revıew*, *12*(4), 487–495.

Pacheco-Cortés, A. M., & Infante-Moro, A. (2020). ICT resignification in a virtual learning environment [La resignificación de las TIC en un ambiente virtual de aprendizaje]. *Campus Virtuales*, *9*(1), 85–99.

Padhan, H., Padhang, P. C., Tiwari, A. V., Ahmed, R., & Hammoudeh, S. (2020). Renewable energy consumption and robust globalization(s) in OECD countries: Do oil, carbon emissions and economic activity matter? *Energy Strategy Reviews*, *32*, 100535. doi:10.1016/j.esr.2020.100535

Pallot, M., & Pawar, K. (2012). A holistic model of user experience for living lab experiential design. In *2012 18th International ICE Conference on Engineering, Technology and Innovation* (pp. 1–15). IEEE. 10.1109/ICE.2012.6297648

Palm, M. (2016). *Technologies of consumer labor: a history of self-service*. Routledge.

Pan, Y., Oksavik, A., & Hildre, H. P. (2020). *Simulator as a Tool for the Future Maritime Education and Research: A Discussion*. Academic Press.

Pandian, A. P. (2019). Artificial intelligence application in smart warehousing environment for automated logistics. *Journal of Artificial Intelligence*, *1*(02), 63–72.

Pangrazio, L., & Sefton-Green, J. (2021). Digital Rights, Digital Citizenship and Digital Literacy: What's the Difference? *Journal of New Approaches in Educational Research*, *10*(1), 15–27. doi:10.7821/naer.2021.1.616

Pantano, E., & Pizzi, G. (2020). Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis. *Journal of Retailing and Consumer Services*, *55*(102096), 1–9. doi:10.1016/j.jretconser.2020.102096

Paoli, C., Voyant, C., Muselli, M., & Nivet, M. L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, *84*(12), 2146–2160. doi:10.1016/j.solener.2010.08.011

Papageorgiou, E. I., Markinos, A. T., & Gemtos, T. (2009). Application of fuzzy cognitive maps for cotton yield management in precision farming. *Expert Systems with Applications*, *36*(10), 12399–12413. doi:10.1016/j.eswa.2009.04.046

Papageorgiou, E. I., & Salmerón, J. L. (2013). A Review of Fuzzy Cognitive Maps Research During the Last Decade. *IEEE Transactions on Fuzzy Systems*, *21*(1), 66–79. doi:10.1109/TFUZZ.2012.2201727

Pappas, N., Caputo, A., Pellegrini, M. M., Marzi, G., & Michopoulou, E. (2021). The complexity of decision-making processes and IoT adoption in accommodation SMEs. *Journal of Business Research*, *131*, 573–583. doi:10.1016/j.jbusres.2021.01.010

Parikh, A., Behnke, C., Vorvoreanu, M., Almanza, B., & Nelson, D. (2014). Motives for reading and articulating user-generated restaurant reviews on Yelp. com. *Journal of Hospitality and Tourism Technology*, *5*(2), 160–176. doi:10.1108/JHTT-04-2013-0011

Park, C. W., Jaworski, B. J., & MacInnis, D. J. (1986). Strategic brand concept-image management. *Journal of Marketing*, *50*(4), 135–145. doi:10.1177/002224298605000401

Parody, C. E. (2011). *A Theory of the Transmedia Franchise Character* (PhD thesis). University of Liverpool.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Parola, F., Satta, G., Notteboom, T., & Persico, L. (2020). Revisiting traffic forecasting by port authorities in the context of port planning and development. *Maritime Economics & Logistics*. Advance online publication. doi:10.105741278-020-00170-7

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1-2), 185–198. doi:10.1016/S1071-5819(03)00017-X

Paschen, J., Kietzmann, J., & Kietzmann, T. C. (2019). Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. *Journal of Business and Industrial Marketing*, *34*(7), 1410–1479. doi:10.1108/JBIM-10-2018-0295

Pata, U. K. (2018). The influence of coal and noncarbohydrate energy consumption on CO2 emissions: Revisiting the environmental Kuznets curve hypothesis for Turkey. *Energy*, *160*, 1115–1123. doi:10.1016/j.energy.2018.07.095

Patten, D. M. (1991). Exposure, legitimacy, and social disclosure. *Journal of Accounting and Public Policy*, *10*(4), 297–308. doi:10.1016/0278-4254(91)90003-3

Peeters, M. M., van Diggelen, J., Van Den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2020). Hybrid collective intelligence in a human–AI society. *AI & Society*, 1–22.

Pelegrín-Borondo, J., Arias-Oliva, M., & Olarte-Pascual, C. (2017). Emotions, price and quality expectations in hotel services. *Journal of Vacation Marketing*, *23*(4), 322–338.

Pelgrín Borondo, J. (2013) *Análisis comparativo de la estrategia de los grupos competitivos: e-clientes vs. clientes off-line* (Doctoral thesis). Universitat Rovira I Virgili, Tarragona.

Pell, G. (2005). Use and misuse of Likert scales. *Medical Education*, *39*(9), 970–970. doi:10.1111/j.1365-2929.2005.02237.x PMID:16150039

Peng, Z., Shan, W., Guan, F., & Yu, B. (2016). Stable vessel-cargo matching in dry bulk shipping market with price game mechanism. *Transportation Research Part E, Logistics and Transportation Review*, *95*, 76–94. doi:10.1016/j.tre.2016.08.007

Perakis, A. N., & Bremer, W. M. (1992). An operational tanker scheduling optimization system: Background, current practice and model formulation. *Maritime Policy & Management*, *19*(3), 177–187. doi:10.1080/751248659

Pérez-Escoda, A., Castro-Subizarreta, A., & Fandos-Igado, M. (2016). The digital competence of Generation Z: keys for its curricular introduction in Primary School. *Scientific Journal of Educommunication, 49*, 71-80. Retrieved from https://www.redalyc.org/pdf/158/15847434008.pdf

Pérez-Rave, I. I., & Muñoz-Giraldo, L. (2014, July - September). Espacio literario relevante sobre la evaluación de la calidad del servicio: Países de realización de los estudios, métodos de análisis, índices de fiabilidad, hipótesis y desafíos. *Elsevier, Ingeniería. Investigación y Tecnología*, *15*(3). Advance online publication. doi:10.1016/S1405-7743(14)70356-7

Perspectives on Issues in AI Governance. (2019). *Google*. https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf.

Peru Retail. (2018). *Peru: Traditional vs. modern channel.* Lima. Retrieved from https://www.peru-retail.com/peru-canal-tradicional-vs-canal-moderno/

Peru Retail. (2020) *What are the most visited shopping channels by Peruvians before Covid-19?* Retrieved from https://www.peru-retail.com/cuales-son-los-canales-de-compra-mas-visitados-por-los-peruanos-ante-el-covid-19/

Perunovic, Z., & Vidic, J.-P. (2011). Innovation in the Maritime Industry. *Proceedings of the 22nd POMS Annual Conferences*.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Petersen, S. (2007). The ethics of robot servitude. *Journal of Experimental & Theoretical Artificial Intelligence*, *19*(1), 43–54. doi:10.1080/09528130601116139

Pfeiffer, J., Pfeiffer, T., Meißner, M., & Weiß, E. (2020). Eye-Tracking-Based Classification of Information Search Behavior Using Machine Learning: Evidence from Experiments in Physical Shops and Virtual Reality Shopping Environments. *Information Systems Research*, *3*(3), 675–691. doi:10.1287/isre.2019.0907

Phelps, J. E., D'Souza, G., & Nowak, G. J. (2001). Antecedents and consequences of consumer privacy concerns: An empirical investigation. *Journal of Interactive Marketing*, *15*(4), 2–17. doi:10.1002/dir.1019

Phelps, J. E., Nowak, G. J., & Ferrell, E. (2000). Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing*, *19*(1), 27–41. doi:10.1509/jppm.19.1.27.16941

Phillip Kottak, C. (2011). Cultural anthropology. McG.

Phillips-Wren, G. (2012). AI tools in decision making support systems: A review. *International Journal of Artificial Intelligence Tools*, *21*(2), 1240005. doi:10.1142/S0218213012400052

Piddle, J. (2017). *Augmented Reality. Where We Will Live*. Springer. doi:10.1007/978-3-319-54502-8

Pilkington, M. (2016). Blockchain technology: principles and applications. In *Research handbook on digital transformations*. Edward Elgar Publishing.

Piller, C. (1999). Everyone is a critic in cyberspace. *Los Angeles Times, 3*(12), A1.

Pine, J., & Gilmore, J. (1999). *The Experience Economy*. Harvard Business School Press.

Pinhasi, A. (2020). *Deploying machine learning models to production: Inference service architecture patterns*. Retrieved from https://medium.com/data-for-ai/deploying-machine-learning-models-to-production-inference-service-architecture-patterns-bc8051f70080

Pirili, M. U. (2011). The Role of Public Investments in Regional Development: A Theoretical Review. *Ege Academic Review*, *11*(2), 309–324.

Plangger, K., & Watson, R. T. (2015). Balancing customer privacy, secrets, and surveillance: Insights and management. *Business Horizons*, *58*(6), 625–633. doi:10.1016/j.bushor.2015.06.006

Plassmann, H., Ramsoy, T. Z., & Milosavljevic, M. (2015). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology*, *22*(1), 18–36. doi:10.1016/j.jcps.2011.11.010

Polyakova, A. G., Loginov, M. P., Serebrennikova, A. I., & Thalassinos, E. I. (2019). Design of a socio-economic processes monitoring system based on network analysis and big data. *International Journal of Economics and Business Administration*, *7*(1), 130–139. doi:10.35808/ijeba/200

Polzin, F., Migendt, M., Taube, ¨. F. A., & von Flotow, P. (2015). Public policy influence on renewable energy investments—A panel data study across OECD countries. *Energy Policy*, *80*, 98–111. doi:10.1016/j.enpol.2015.01.026

Poole, D. L., & Mackworth, A. K. (2010). *Artificial intelligence: Foundations of computational agents*. Cambridge University Press. doi:10.1017/CBO9780511794797

Pope, R. I. (2003). *Kubrick's Crypt, a Derrida/Deluze Monster, in 2001: A Space Odyssey*. Edinburgh University Press Journals.

Port of Antwerp deploys autonomous drones for safety enforcement. (2021, February). *Dry Cargo International*. Retrieved from https://www.drycargomag.com/port-of-antwerp-deploys-autonomous-drones-for-safety-enforcement

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Portugal.EU. (2021). *Program for the Portuguese Presidency of the Council of the European Union*. Author.

Portuguese Government Portal. (2021). *Americas - Regional Affairs - Foreign Policy - Diplomatic Portal. Ministry of Foreign Affairs of Portugal*. Retrieved from https://www.portaldiplomatico.mne.gov.pt/en/foreign-policy/regional-affairs/americas

Porwik, P. (2010). *The Modern Techniques of Latent Fingerprint Imaging*. IEEE. doi:10.1109/CISIM.2010.5643695

Poulis, K., Galanakis, G. C., Triantafillou, G. T., & Poulis, E. (2020). Value migration: Digitalization of shipping as a mechanism of industry dethronement. *Journal of Shipping and Trade*, *5*(1), 1–18. doi:10.118641072-020-00064-0

Pradeep, A. K., Appel, A., & Sthanunathan, S. (2019). *AI for Marketing and Product Innovation*. Wiley.

Prahono, A., & Elidjen. (2015). Evaluating the role e-government on public administration reform: Case of official city government websites in indonesia. *Procedia Computer Science*, *59*, 27–33. doi:10.1016/j.procs.2015.07.334

Prassas, G., Pramatari, K. C., Papaemmanouil, O., & Doukidis, G. J. (2001). A recommender system for online shopping based on past customer behavior. *Proceedings of the 14th Bled Electronic Commerce Conference e-Everything: e-Commerce, e-Government, e-Household, e-Democracy*.

Prebensen, N., Kim, H., & Uysal, M. (2015). Cocreation as Moderator between the Experience Value and Satisfaction Relationship. *Journal of Travel Research*, 1–12.

Premkumar, G., Ramamurthy, K., & Crum, M. (1997). Determinants of EDI adoption in the transportation industry. *European Journal of Information Systems*, *6*(2), 107–121. doi:10.1057/palgrave.ejis.3000260

Premkumar, G., & Roberts, M. (1999). Adoption of new information technologies in rural small businesses. *Omega*, *27*(4), 467–484. doi:10.1016/S0305-0483(98)00071-1

Prescott, J., (2016). Why watching Westworld's robots should make us question ourselves. *The Conversation*.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. doi:10.1016/S0001-6918(99)00050-5 PMID:10769936

Proietti, M. (2019). Experimental test of local observer independence. *Science Advances*, *5*(9). Advance online publication. doi:10.1126ciadv.aaw9832

Projeto JUL- LSW contributes to the concept of the National Single Window. (2021). Retrieved from https://www.projeto-jul.pt/en/news/lsw-contributes-to-the-concept-of-national-single-window

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking* (1st ed.). O'Reilly Media.

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

Przychodzen, W., & Przychodzen, J. (2020). Determinants of renewable energy production in transitioneconomies: A panel data approach. *Energy*, *191*, 116583. doi:10.1016/j.energy.2019.116583

Puente-Bienvenido, H. & Sequeiros, C. (2019a). A Sociological Look at Gaming Software: Erving Goffman's Dramaturgy in Video Games. *Revista Española de Investigaciones Sociológicas*, *166*, 135-152. doi:10.5477/cis/reis.166.135

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Puente-Bienvenido, H., & Sequeiros Bruna, C. (2019b) Goffman y los videojuegos: Una aproximación sociológica desde la perspectiva dramatúrgica a los dispositivos videolúdicos. *Revista Española de Sociología*, *28*(2). https://recyt.fecyt.es/index.php/res/article/view/68948

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, *85*(1), 131–151. doi:10.1177/0022242920953847

PuroMarketing. (2019). *En 2030 la Experiencia de Cliente dependerá de la Inteligencia Artificial*. Retrieved December 18, 2020, from https://www.puromarketing.com/12/32895/experiencia-cliente-dependera-inteligencia-artificial.html

PWC. (2017a). *La Inteligencia Artificial impulsará el PIB mundial un 14% en 2030 por sus efectos en la productividad y en el consume*. Retrieved December 18, 2020, from https://www.pwc.es/es/sala-prensa/notas-prensa/2017/la-inteligencia-artificial-impulsara-pib-mundial.html

PWC. (2017b). *Sizing the prize. What's the real value of AI for your business and how can you capitalise?* Retrieved December 8, 2020, from https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

PWC. (2018). *Will robots really steal our jobs? An international analysis of the potential long term impact of automation*. PWC. https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/impact_of_automation_on_jobs.pdf

PWC. (2020). *Experience is everything: Here's how to get it right.* Retrieved December 15, 2020, from https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/pwc-consumer-intelligence-series-customer-experience.pdf

Qian, P., Li, A., & Liu, M. (2019, June). Latent fingerprint enhancement based on DenseUNet. In *2019 International Conference on Biometrics (ICB)* (pp. 1-6). IEEE.

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 67. doi:10.118613634-016-0355-x

Qualtrics. (2020). *What is closed-loop customer experience management?* Retrieved December 16, 2020, from https://www.qualtrics.com/experience-management/customer/closed-loop-cx/

Quéant, P. (2019). IA et puissance aérospatiale, implications pour l'Armée de l'air [AI and Aerospace Power, Implications for the Air Force]. *Revue Défense Nationale*, *820*(5), 117–122. doi:10.3917/rdna.820.0117

Radziwon, A., Bilberg, A., Bogers, M., & Madsen, E. S. (2014). The smart factory: Exploring adaptive and flexible manufacturing solutions. *Procedia Engineering*, *69*, 1184–1190. doi:10.1016/j.proeng.2014.03.108

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. doi:10.100711747-019-00710-5

Raine, R. (2009). Making A Clever Intelligent Agent: The Theory Behind The Implementation. *IEEE International Conference on Intelligent Computing and Intelligent Systems*. 10.1109/ICICISYS.2009.5358137

Raman, A., & Hoder, C. (2020). *Building intelligent apps with cognitive APIs*. O'Reilly.

Ramírez, M.-S., & García-Peñalvo, F.-J. (2018). Co-Creation and Open Innovation: Systematic Literature Review. *Comunicar: Media Education Research Journal*, *26*(54), 9–18. doi:10.3916/C54-2018-01

Ramos, K., Cuamea, O., Morgan, J., & Estrada, A. (2020). Social Networks' Factors Driving Consumer Restaurant Choice: *An Exploratory Analysis. In International Conference on Applied Human Factors and Ergonomics* (pp. 158-164). Springer.

Ranchhod, A. (2007). Developing mobile marketing strategies. *International Journal of Mobile Advertising*, *2*(1), 76–83.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Randhawa, K., & Scerri, M. (2015). Service innovation: A review of the literature. In *The handbook of service innovation* (pp. 27–51). Springer. doi:10.1007/978-1-4471-6590-3_2

Real Academia Española, R. A. E. (2019). *Diccionario de la lengua española* (22nd ed.). Retrieved December 17, 2020, from http://www.rae.es

Recio Muñoz, F., Silva Quiroz, J., & Abricot Marchant, N. (2020). Análisis de la Competencia Digital en la Formación Inicial de estudiantes universitarios: Un estudio de meta-análisis en la Web of Science. *Píxel-Bit, Revista De Medios Y Educación*, *59*, 125-146. [] doi:10.12795/pixelbit.77759

Regional Government of Arequipa. (2020). *Regional Executive Resolution N° 192-2020-GRA/GR.* Retrieved from www.regionarequipa.gob.pe: https://www.regionarequipa.gob.pe/Cms_Data/Contents/GobRegionalArequipaInv/Media/Resolucion.Detalle/2020/R.E.R/RER-192-2020.pdf

Reichheld, F. (2014). How the Net Promoter Score (NPS) Can Drive Growth. *JWI 518: Marketing in The Global Environment*, 73-100.

Resıdency of the Republıc of Turkey Strategy and Budget Dırectorate (SBB). (n.d.). Retrieved from http: http://www.sbb.gov.tr/wp-content/uploads/2019/04/Yat%C4%B1r%C4%B1mlar%C4%B1n-%C4%B0llere-G%C3%B6re-Da%C4%9F%C4%B1l%C4%B1m%C4%B1.xlsx

Revathy, T., Pramila, G., Adhiraja, A., & Askerunisa, A. (2014, April). Automatic Latent Fingerprint Segmentation based on Orientation and Frequency Features. In *2014 International Conference on Communication and Signal Processing* (pp. 1192-1196). IEEE. 10.1109/ICCSP.2014.6950029

Revieve. (n.d.). *AI Skincare Advisor*. https://www.revieve.com/solutions/skincareadvisor

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). *"Why should I trust you?": Explaining the predictions of any classifier.* Paper presented at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA. 10.1145/2939672.2939778

Richardson, K. (2016). Sex robot matters: Slavery, the prostituted, and the rights of machines. *IEEE Technology and Society Magazine*, *35*(2), 46–53. doi:10.1109/MTS.2016.2554421

Richardson, S. (2020). Cognitive automation: A new era of knowledge work? *Business Information Review*, *37*(4), 182–189. doi:10.1177/0266382120974601

Rico, L., & Ohlenschläger, K. (2005). *Banquete 05 – Communication in Evolution*. Museo Conde Duque.

Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, *29*, 37–58. doi:10.1016/j.accinf.2018.03.001

Riley, D., Charlton, N., & Wason, H. (2015). The impact of brand image fit on attitude towards a brand alliance. *Management & Marketing*, *10*(4), 270–283. doi:10.1515/mmcks-2015-0018

Rivera-Castro, R., Pilyugina, P., Pletnev, A., Maksimov, I., Wyz, W., & Burnaev, E. (2019). Topological data analysis of time series data for B2B Customer Relationship Management. https://arxiv.org/abs/1906.03956

Rivoir, A., Morales, M. J., & Casamayou, A. (July 18, 2019). Uses and perceptions of digital technologies in older people. Limitations and benefits for their quality of life. *Austral Journal of Social Sciences*, 295-313. doi:10.4206/rev.austral.cienc.soc.2019.n36-15

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Robinson, S., Orsingher, C., Alkire, L., De Keyser, A., Giebelhausen, M., Papamichail, K. N., Shams, P., & Temerak, M. S. (2020). Frontline encounters of the AI kind: An evolved service encounter framework. *Journal of Business Research*, *116*, 366–376. doi:10.1016/j.jbusres.2019.08.038

Rodrigue, J.-P. (2010). *Maritime Transportation: Drivers for the Shipping and Port Industries*. Academic Press.

Rodrigues, K. R., Ferreira, C. G., Murta, A. L. S., & Murta, M. P. A. (2017). Sistema Portuário Brasileiro e o Uso da Tecnologia para uma Gestão Eficiente. *HOLOS*, (7), 110–126. doi:10.15628/holos.2017.6080

Rodriguez Rojas, C. I. (2015). *Sostenibilidad en las empresas*. Recuperado 6 de julio de 2020, de Escuela de Organización InduStrial website: https://www.eoi.es/blogs/carollirenerodriguez/2012/05/20/sostenibilidad-en-las-empresas/

Rodríguez-Ardura, I., Maraver-Tarifa, G., Jiménez-Zarco, A. I., Martínez-Argüelles, M. J., & Ammetller, G. (2018). *Principios y estrategias de marketing* (Vol. 2). Edi UOC.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *The American Psychologist*, *42*, 873. doi:10.1037/h0092053 PMID:2221571

Rogers Everett, M. (1995). *Diffusion of innovations*.

Rogers, E. (2003). Diffusion of innovations. Delran, NJ: Simon & Schuster.

Roggeveen, A. L., Tsiros, M., & Grewal, D. (2012). Understanding the co-creation effect: When does collaborating with customers provide a lift to service recovery? *Journal of the Academy of Marketing Science*, *40*(6), 771–790. doi:10.100711747-011-0274-1

Rohit, Dharamshi, & Subramanyam. (2019). Approaches to Question Answering Using LSTM and Memory Networks: SocProS 2017. In *Soft Computing for Problem Solving*. Doi:10.1007/978-981-13-1592-3 15

Rohit, G. (2019). *Dharamshi, Ekta Subramanyam, Natarajan*. Approaches to Question Answering Using LSTM and Memory.

Rollins, J. (2015). *Why we need a methodology for data science*. Retrieved from https://www.ibmbigdatahub.com/blog/why-we-need-methodology-data-science

Romano, A. (2020, July 5). This hotel has a robot named Rosé that will deliver wine to your room Without human contact. *Travel and Leisure*. https://www.travelandleisure.com/food-drink/wine/rose-robot-room-service-hotel-trio-california

Romero, J., Ruiz-Equihua, D., Loureiro, S. M. C., & Casaló, L. V. (2021). *Smart Speaker Recommendations: Impact of Gender Congruence and Amount of Information on Users' Engagement and Choice*. Frontier in Psychology., doi:10.3389/fpsyg.2021.659994

Ronen, D. (1993, December). Ship scheduling: The last decade. *Ship scheduling: The last decade, 71*(3), 325-333. doi:10.1016/0377-2217(93)90343-L

Ronen, D. (1983, February). Cargo ships routing and scheduling: Survey of models and problems. *European Journal of Operational Research*, *12*(2), 119–126. doi:10.1016/0377-2217(83)90215-1

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, *9*(1), 86–136. doi:10.1177/1536867X0900900106

Rorissa, A., & Demissie, D. (2010). An analysis of african e-government service websites. *Government Information Quarterly*, *27*(2), 161–169. doi:10.1016/j.giq.2009.12.003

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rosenblatt, F. (1957). The perceptron - A perceiving and recognizing automaton. *Issues (Chicago, Ill.)*, *Technical Report*, 85–460–1.

Rossi, F. (2018). Building trust in artificial intelligence. *Journal of International Affairs*, *72*(1), 127–134.

Rossiter, J. R., Silberstein, R. B., Harris, P. G., & Nield, G. (2001). So What? A Rejoinder to the Reply by Crites and Aikman-Eckenrode to Rossiter et al. (2001). *Journal of Advertising Research*, *41*(3), 59–61. Advance online publication. doi:10.2501/JAR-41-3-59-61

Ross, J. (2018). The fundamental flaw in AI implementation. *MIT Sloan Management Review*, *59*(2), 10–11.

Roszkowski, M., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, *35*(1), 117–134. doi:10.1080/02602930802618344

Roth, C., & Koenitz, H. (2016). Evaluating the User Experience of Interactive Digital Narrative. In *Proceedings of the 1st International Workshop on Multimedia Alternate Realities*. Association for Computing Machinery. 10.1145/2983298.2983302

Rouf, M. A. (2011). The Corporate Social Responsibility Disclosure: A Study of Listed Companies in Bangladesh. *Business and Economics Research Journal*, *2*(3), 19–32.

Ruiz, J. B., & Bell, R. A. (2021). Predictors of intention to vaccinate against COVID-19: Results of a nationwide survey. *Vaccine*, *39*(7), 1080–1086. doi:10.1016/j.vaccine.2021.01.010 PMID:33461833

Rusk, N. (2016). Deep learning. *Nature Methods*, *13*(1), 35–35. doi:10.1038/nmeth.3707 PMID:27110626

Russel, S. (2015). Ethics of Artificial Intelligence. *Nature*, 415-418.

Russell, J. A. (2003). Core affect and psychological construction of emotion. *Psychological Review*, *110*(1), 145–172.

Russell, J. A., Lewicka, M., & Niit, T. (1989). A Cross-Cultural Study of a Circumflex Model of Affect. *Journal of Personality and Social Psychology*, *57*(5), 848–856. doi:10.1037/0022-3514.57.5.848

Russell, S. J. (1997). Rationality and intelligence. *Artificial Intelligence*, *94*(1-2), 57–77. doi:10.1016/S0004-3702(97)00026-X

Russell, S. J., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson.

Russkikh, P., & Kapulin, D. (2020). Simulation modeling for optimal production planning using Tecnomatix software. *Journal of Physics: Conference Series*, *1661*, 012188. doi:10.1088/1742-6596/1661/1/012188

Russo, S. J. (2020). Is de-identification of personal health information in the age of artificial intelligence a reality or a noble myth? *Journal of Health Care Compliance*, (March–April), 55–59.

Rust, R. T. (2020). The future of marketing. *International Journal of Research in Marketing*, *37*(1), 15–26. doi:10.1016/j.ijresmar.2019.08.002

Rust, R. T., & Huang, M. H. (2014). The service revolution and the transformation of marketing science. *Marketing Science*, *33*(2), 206–221. doi:10.1287/mksc.2013.0836

Ryan, M. L. (2001). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. Johns Hopkins University Press.

Sabhanayagam, T., Venkatesan, V. P., & Senthamaraikannan, K. (2018). A comprehensive survey on various biometric systems. *International Journal of Applied Engineering Research: IJAER*, *13*(5), 2276–2297.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sabio & DEC. (2020). *La omnicanalidad es responder a lo que el cliente te está pidiendo*. https://brilliantcx.sabiogroup.com/l/668023/2020-09-30/25ly6

Salas-Rueda, R.-A., & Salas-Silis, J.-A. (2019). WeVideo: ¿Servicio en la nube útil para los estudiantes durante la construcción y presentación de los contenidos audiovisuales? *Vivat Academia*, *1*, 67–89. doi:10.15178/va.2019.149.67-89

Salazar, M. S. (2015). The dilemma of combining positive and negative items in scales. *Psicothema*, *27*(2), 192–199. PMID:25927700

Salehi-Esfahani, S., Ravichandran, S., Israeli, A., & Bolden, E. III. (2016). Investigating information adoption tendencies based on restaurants' user-generated content utilizing a modified information adoption model. *Journal of Hospitality Marketing & Management*, *25*(8), 925–953. doi:10.1080/19368623.2016.1171190

Sallam, M. (2021). COVID-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates. *Vaccines*, *9*(2), 160. doi:10.3390/vaccines9020160 PMID:33669441

Same, S., & Larimo, J. (2012, May). Marketing theory: Experience marketing and experiential marketing. In *7th International Scientific Conference. Business and Management* (pp. 10-11). Academic Press.

Sangthong, M. (2020). The Effect of the Likert Point Scale and Sample Size on the Efficiency of Parametric and Nonparametric Tests. *Thailand Statistician*, *18*(1), 55–64.

Sanmartín, A. y Megías, I. (2020). *Jóvenes, futuro y expectativa tecnológica*. Madrid: Centro Reina Sofía sobre Adolescencia y Juventud, Fad. Doi:10.5281/zenodo.3629108

Sanny, L., Arina, A., Maulidya, R., & Pertiwi, R. (2020). Purchase intention on Indonesia male's skin care by social media marketing effect towards brand image and brand trust. *Management Science Letters*, *10*(10), 2139–2146.

Santanen, E. (2019). The value of protecting privacy. *Business Horizons*, *62*(1), 5–14. doi:10.1016/j.bushor.2018.04.004

Santolaria, C. J. S., & González, J. P. G. (2013). Las memorias de sostenibilidad y su divulgación. *Contaduría Universidad de Antioquia*, (57), 107-118. Retrieved from https://revistas.udea.edu.co/index.php/cont/article/view/15580/13523

Sarenet (2019). *2022: Más tráfico en Internet que el acumulado en tres décadas.* https://blog.sarenet.es/trafico-web/

Sarkar, D. (2018). *Get Smarter with Data Science — Tackling Real Enterprise Challenges. Take your Data Science Projects from Zero to Production.* Towards Data Science. Accessed and reviewed in April 2021, https://towardsdatascience.com/get-smarter-with-data-science-tackling-real-enterprise-challenges-67ee001f6097

SAS. (2021). *History of Big Data*. SAS. Accessed and reviewed in April 2021, https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Sasmaz, M. U., Sakar, E., Yayla, Y. E., & Akkucuk, U. (2020). The Relationship between Renewable Energy and Human Development in OECD Countries: A Panel Data Analysis. *Sustainability*, *12*(18), 7450. doi:10.3390u12187450

Sato, D., Wider, A., & Windheuser, C. (2019). *Continuous delivery for machine learning: Automating the end-to-end lifecycle of machine learning applications*. Retrieved from https://martinfowler.com/articles/cd4ml.html

Saur-Amaral, I., Ferreira, P., & Conde, R. (2013). Linking past and future research in tourism management through the lens of marketing and consumption: a systematic literature review Investigação passada e futura sobre gestão do turismo, marketing e consumo: uma revisão sistemática da literatura. *Conde / Tourism & Management Studies, 9*(1), 35–40.

Saxena, D., & Lamest, M. (2018). Information overload and coping strategies in the big data context: Evidence from the hospitality sector. *Journal of Information Science*, *44*(3), 287–297. doi:10.1177/0165551517693712

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Scarsi, R. (2007). The bulk shipping business: Market cycles and shipowners' biases. *Maritime Policy & Management*, *34*(6), 577–590. doi:10.1080/03088830701695305

Schaar, P. (2010). Privacy by design. *Identity in the Information Society*, *3*(2), 267–274. doi:10.100712394-010-0055-x

Schell, J. (2008). The Art of Game Design: A book of lenses. *Elsevier CRC Press*.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Sciences Information. Information Sur les Sciences Sociales*, *44*(4), 695–729.

Schim van der Loeff, W., Godar, J., & Prakasha, V. (2018, December). A spatially explicit data-driven approach to calculating commodity-specific shipping emissions per vessel. *Journal of Cleaner Production*, *205*, 895–908. doi:10.1016/j.jclepro.2018.09.053

Schmidt, L., Skvortsova, V., Kullen, C., Weber, B., & Plassmann, H. (2017). How context alters value: The brain's valuation and affective regulation system link price cues to experienced taste pleasantness. *Scientific Reports*, *7*(1), 8098. Advance online publication. doi:10.103841598-017-08080-0 PMID:28808246

Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management*, *15*(1-3), 53–67.

Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement*, *41*(4), 1101–1114. doi:10.1177/001316448104100420

Schuller, D., & Schuller, B. W. (2018). The age of artificial emotional intelligence. *Computer*, *51*(9), 38–46. doi:10.1109/MC.2018.3620963

Schuurman, D., De Marez, L., & Ballon, P. (2015). Living Labs: a systematic literature review. Open Living Lab Days 2015.

Schwab, K. (2017). *The fourth industrial revolution*. Currency.

Schwartz, L. (2019). *How to Use Case Studies in Your Classroom to Make Learning Relevant*. Edutopia.

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management*, *35*(7-8), 693–715. doi:10.1080/0267257X.2019.1596970

Sci-Kit Learn. (2020). *API Reference*. https://scikit-learn.org/stable/modules/classes.html#module-sklearn.preprocessing

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., & Young, M. (2014). *Machine learning: The high interest credit card of technical debt*. Paper presented at the SE4ML Workshop (NIPS 2014).

Semrl, J., & Matei, A. (2017). Churn prediction model for effective gym customer retention. *Proceedings of 4th International Conference on Behavioral, Economic, and Socio-Cultural Computing, BESC 2017,* 1–3. 10.1109/BESC.2017.8256385

Sequeiros, C., & Puente-Bienvenido, H. (2018). Sociología y ci-fi: repensando los procesos de cambio social desde la ciencia ficción. *Nudos. Sociología, Teoría y didáctica de la literature, 2*, 4-16. doi:10.24197/nrtstdl.1.2018.4-16

Sequeiros, C., & Puente-Bienvenido, H. (2019). Cambio social, tecnología y ciencia-ficción. *Sociología y Tecnociencia*, *9*(2), 115–138.

Serban, A., van der Blom, K., Hoos, H., & Visser, J. (2020). *Adoption and effects of software engineering best practices in machine learning*. Paper presented at the ESEM conference. 10.1145/3382494.3410681

Serrano, C. L., Peña Lapeira, C. J., & Laverde Guzmán, M. Y. (2018). Influence of ICT in the economic development of Colombia. *Revista Ciencias de la Información*, 3-10. Retrieved from http://cinfo.idict.cu/index.php/cinfo/article/view/854

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sestino, A., Prete, M. I., Piper, L., & Guido, G. (2020). Internet of Things and Big Data as enablers for business digitalization strategies. *Technovation*, *98*, 102173. doi:10.1016/j.technovation.2020.102173

Shankar, V., Smith, A. K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, *20*(2), 153–175.

Shelley, M. (1818). *Frankenstein or The Modern Prometheus*. Harding, Mavor and Jones.

Shen, B., Choi, T.-M., & Chow, P.-S. (2017). Brand loyalties in designer luxury and fast fashion co-branding alliances. *Journal of Business Research*, *81*, 173–180. doi:10.1016/j.jbusres.2017.06.017

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*, 97–112. doi:10.1016/j.ijpe.2017.06.006

Sherstins, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. *Physica D: Nonlinear Phenomena, 404*.

Sheth, J. (2020). Impact of Covid-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research 117*, 280-283. doi:10.1016/j.jbusres.2020.05.059

Shiau, W. L., Sarstedt, M., & Hair, J. F. (2019). Internet research using partial least squares structural equation modeling (PLS-SEM). *Internet Research*.

Shimp, T., & Sharma, S. (1987). Consumer ethnocentrism: Construction and validation of the CETSCALE. *JMR, Journal of Marketing Research*, *26*(August), 280–289. doi:10.1177/002224378702400304

Shin, D. (2010). The effects of trust, security & privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, *22*(5), 428–438. doi:10.1016/j.intcom.2010.05.001

Shipwright design and construction, additional design procedures, design code for unmanned marine systems. (n.d.). *Lloyd's Register*. Retrieved from www.lr.org

Shrestha, Y. R., Krishna, V., & von Krogh, G. (2020). Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *Journal of Business Research*, *123*, 588–603. doi:10.1016/j.jbusres.2020.09.068

Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). A Comparison of ARIMA and LSTM in Forecasting Time Series. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 1394–1401. 10.1109/ICMLA.2018.00227

Siegel, S. (1957). Nonparametric statistics. *The American Statistician*, *11*(3), 13–19.

Sierra Diez, B., Froufe Torres, M., & Falces Delgado, C. (2010). El papel de las metas conscientes e inconscientes en la motivación del consumidor. *REME*, *13*(35-36), 193–210.

Siggelkow, N. (2007). Persuasion with case studies. *Academy of Management Journal*, *50*(1), 20–24. doi:10.5465/amj.2007.24160882

Sikander, M., & Khiyal, H. (2018). *Computational Models for Upgrading Traditional Agriculture*. Preston University. Accessed and reviewed in April 2021, https://www.researchgate.net/publication/326080886

Sikdar, S. (2018). Artificial Intelligence, its impact on innovation, and the Google effect. *Clean Technologies and Environmental Policy*, *20*, No-1–No-2.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Silverstein, S. (2018). REX real estate: Home showings-AI / Voice / Chatbot. *Behance*. https://www.behance.net/gallery/64004179/REX-Real-Estate-Home-Showings-AI-Voice-Chatbot

Simonin, J., Bertin, E., Traon, Y., Jezequel, J.-M., & Crespi, N. (2010). Business and Information System Alignment: A Formal Solution for Telecom Services. In *2010 Fifth International Conference on Software Engineering Advances*. IEEE. 10.1109/ICSEA.2010.49

Singer, E. (2010). The slow rise of the robot surgeon. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/news/418141/the-slow-rise-of-the-robot-surgeon/

Singh, J. P., Kalafatis, S., & Ledden, L. (2014). Consumer perceptions of cobrands: The role of brand positioning strategies. *Marketing Intelligence & Planning*, *32*(2), 145–159. doi:10.1108/MIP-03-2013-0055

Singler, E. (2015). *Green Nudge: Changer les comportements pour sauver la planète*. Pearson.

Skålén, P., Gummerus, J., Von Koskull, C., & Magnusson, P. R. (2015). Exploring value propositions and service innovation: A service-dominant logic study. *Journal of the Academy of Marketing Science*, *43*, 137–158. doi:10.100711747-013-0365-2

Slocombe, W. (2016). What science fiction tells us about our trouble with Artificial Intelligence, *The Conversation*.

Sloman, A. (2006). *Why Asimov's three laws of robotics are unethical.* Retrieved June 9, 2006, URL: https://www.cs.bham.ac.uk/research/projects/cogaff/misc/asimov-three-laws.html

Smarandescu, L., & Shimp, T. A. (2015). Drink Coca-Cola, eat popcorn, and choose powerade: Testing the limits of subliminal persuasion. *Marketing Letters*, *26*(4), 715–726. doi:10.100711002-014-9294-1

SmartPort. (2019). Smart ships and the changing maritime ecosystem (T. van Dijk, H. van Dorsser, R. van den Berg, H. Moonen, & R. Negenborn, Eds.). Academic Press.

Smith, J. H., Milberg, S. J., & Burke, J. B. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *Management Information Systems Quarterly*, *20*(2), 167–196. doi:10.2307/249477

Snow, J. (2018, July 26). Amazon's face recognition falsely matched 28 members of congress with mugshots. *ACLU*. https:// www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Sofi, S. A., Nika, F. A., Shah, M. S., & Zarger, A. S. (2018). Impact of subliminal advertising on consumer buying behaviour: An empirical study on young Indian consumers. *Global Business Review*, *19*(6), 1580–1601. doi:10.1177/0972150918791378

Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics*, *47*, 101324. Advance online publication. doi:10.1016/j.tele.2019.101324

Solana Gutierrez, J., Rincón Sanz, G., Alonso González, C., & Garcia De Jalon Lastra, D. (2015). Use of Maps of Diffuse Knowledge (FCMs) in the prioritization of river restoration: Application to the Esla River [Utilización de Mapas de Conocimiento Difuso (MCD) en la asignación de prioridades de la restauración fluvial: Aplicación al río Esla]. *Cuadernos de la Sociedad Española de Ciencias Forestales*, *41*, 367–380.

Soler, A. S., Ort, M., Steckel, J., & Nieschullze, J. (2016). *An introduction to data management*. Accessed from: https://www.gfbio.org/documents/10184/22817/Reader_GFBio_BefMate_20160222/1ca43f24-255044b3-a05e-e180c3e544c0

SolomonR. C. (2017). *Emotion*. Retrieved from http://www.britannica.com/topic/emotion

Sommer, L. (2015). Industrial revolution-industry 4.0: Are German manufacturing SMEs the first victims of this revolution? *Journal of Industrial Engineering and Management*, *8*(5), 1512–1532. doi:10.3926/jiem.1470

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sommerville, I. (2019). *Artificial intelligence and systems engineering*. Academic Press.

Sommerville, I., & Sawyer, P. (1997). *Requirements engineering: A good practice guide*. John Wiley & Sons, Inc.

Song, I.-Y., & Zhu, Y. (2017). Big Data and Data Science: Opportunities and Challenges of iSchools. *Journal of Data and Information Science*, *2*(3), 1–18. doi:10.1515/jdis-2017-0011

Sonnenwald, D. H., Maglaughlin, K. L., & Whitton, M. C. (2001). Using innovation diffusion theory to guide collaboration technology evaluation: work in progress. In *Proceedings Tenth IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises. WET ICE 2001* (pp. 114-119). IEEE. 10.1109/ENABL.2001.953399

Sorrel, C. (2016, November 13). Self-driving Mercedes will be programmed to sacrifice pedestrians to save the driver. *Fast Company*. https://www.fastcompany.com/3064539/self-driving-mercedes-will-be-programmed-to-sacrifice-pedestrians-to-save-the-driver

Sosa Sierra, M. D. C. (2011). Artificial intelligence in business financial management [Inteligencia artificial en la gestión financiera empresarial]. *Revista científica Pensamiento y gestión*, (23), 153-186.

Spence, C. (2019). Neuroscience-inspired design: From academic neuromarketing to commercially relevant research. *Organizational Research Methods*, *22*(1), 275–298.

Spiekermann, S. (2015). *Ethical IT innovation: A value-based system design approach*. CRC Press. doi:10.1201/b19060

Srour, F., Oosterhout, M., Baalen, P. J., & Zuidwijk, R. (2008). *Port Community System Implementation: Lessons Learned from International Scan*. Academic Press.

Stahl, B. C. (2021). Organisational responses to the ethical issues of AI. *AI & Society*.

Stein, M. (2020, April). *Unmanned maritime infrastructure inspection-a mixed-method risk management approach from German port facilities*. Academic Press.

Stein, M. (2018, June). Integrating Unmanned Vehicles in Port Security Operations: An Introductory Analysis and First Applicable Frameworks. *Ocean Yearbook Online*, *32*(1), 556–583. doi:10.1163/22116001-03201022

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*. *New Series*, *103*(2684), 677–680.

Stoicescu, C. (2016). Big Data, the perfect instrument to study today's consumer behavior. *Database System Journal*, *6*, 28–42.

Stoll, M., Baecke, S., & Kenning, P. (2008). What they see is what they get? An fMRI-study on neural correlates of attractive packaging. *Journal of Consumer Behaviour*, *7*(4-5), 342–359. Advance online publication. doi:10.1002/cb.256

Stopford, M. (2009). *Maritime Economics*. Routledge.

Strotmann, H. (2007). Entrepreneurial survival. *Small Business Economics*, *28*(1), 87–104. doi:10.100711187-005-8859-z

Sukhbaatar, S., & Arthur, S. J. W. R. F. (n.d.). *End-To-End Memory Networks*. Dept. of Computer Science Courant Institute, New York University.

Sullivan, G. M., & Artino, A. R. Jr. (2013). Analyzing and interpreting data from Likert type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. doi:10.4300/JGME-5-4-18 PMID:24454995

Sultana, N., & Nangunoori, R. (2008). Ethical Acceptability of Neuromarketing: Relevance, Limits and Limitations. *Prabandhan: Indian Journal of Management*, *1*(1), 31. Advance online publication. doi:10.17010//2008/v1i1/64654

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sumioka, H., Nakae, A., Kanai, R., & Ishiguro, H. (2013). Huggable communication medium decreases cortisol levels. *Scientific Reports*, *3*(1), 1–6. doi:10.1038rep03034 PMID:24150186

Su, N., & Reynolds, D. (2017). Effects of brand personality dimensions on consumers' perceived self-image congruity and functional congruity with hotel brands. *International Journal of Hospitality Management*, *66*, 1–12. doi:10.1016/j.ijhm.2017.06.006

Sung, C., Zhang, B., Higgins, C. Y., & Choe, Y. (2016). Data-driven sales leads prediction for everything-as-a-service in the cloud. *Proceedings - 3rd IEEE International Conference on Data Science and Advanced Analytics, DSAA 2016*, (May), 557–563. 10.1109/DSAA.2016.83

Sun, S., Wei, Y., Tsui, K., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1–10. doi:10.1016/j.tourman.2018.07.010

Superintencia de Banca. Seguros y AFP - BID. (2012). *Pilot Survey on Access to and Use of Financial Services 2012.* Retrieved from https://www.sbs.gob.pe/Portals/0/jer/est_incl_finan/Informe%20de%20resultados%20inferenciales_.pdf

Susskind, R., & Susskind, D. (2016, October 11). Technology will replace many doctors, lawyers, and other professionals. *Harvard Business Review*. https://hbr.org/2016/10/robots-willreplace-doctors-lawyers-and-other-professionals

Sutherland, I. E. (1965). *The Ultimate Display. Multimedia: From Wagner to Virtual Reality*. Norton.

Svoboda, J., Monti, F., & Bronstein, M. M. (2017, October). Generative convolutional networks for latent fingerprint reconstruction. In *2017 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 429-436). IEEE. 10.1109/BTAS.2017.8272727

Swaid, S. I., & Wigand, R. T. (2012). The effect of perceived site-to-store service quality on perceived value and loyalty intentions in multichannel retailing. *International Journal of Management*, *29*(3), 301.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 1–9. 10.1109/CVPR.2015.7298594

Taddy, M. (2019). *Business Data Science: Combining Machine Learning and Economics to Optimize* (1st ed.). Automate, and Accelerate Business Decisions.

Tamr. (2014). *The Evolution of ETL.* Tamr. http://www.tamr.com/evolution-etl/

Tan, D. (2019). *Coding habits for data scientists*. Retrieved from https://www.thoughtworks.com/insights/blog/coding-habits-data-scientists

Tanev, S., & Sandstrom, G. (2019). Editorial Artificial Intelligence and Innovation Management. *Technology Innovation Management Review*, *9*(12), 3–4. doi:10.22215/timreview/1286

Tat, C. H., Lim, K., Lund, S., Deggim, H., Koyama, T., Foo, D., . . . de Souza, C. (2020, October). *Future of shipping: Digitalisation webinar.* Retrieved from https://fos-digitalisation2020.sg/

Taylor, S. J., & Bodgan, R. (1984). *Introducción. Ir hacia la gente. Introducción a los métodos cualitativos de investigación. La búsqueda de significados*. Paidós Ibérica.

Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, *72*(1), 37–45. doi:10.1080/00031305.2017.1380080

Tecnohotel. (2018). *Artificial Intelligence, advantages and challenges of automating the hotel industry* [Inteligencia Artificial, ventajas y retos de automatizar la industria hotelera]. https://n9.cl/lyia

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tejedor-Flores, N., Galindo-Villardón, P., & Vicente-Galindo, P. (2017). Sustainability multivariate analysis based on the global reporting initiative (GRI) framework, using as a case study: Brazil compared to Spain and Portugal. *International Journal of Sustainable Development and Planning*, *12*(4), 667–677. doi:10.2495/SDP-V12-N4-667-677

Tejedor, N. (2015). *Análisis Multivariante de la Sosteniblidad en Brasil para los años 2011, 2012 y 2013, a través del Global Reporting Initiative (GRI) (Trabajo fin de máster)*. Universidad de Salamanca.

Telecom Advisory Services, LLC. (2015). *The Digital Ecosystem and Economy in Latin America*. Retrieved from http://www.teleadvs.com/wp-content/uploads/Presentacion_IGF_v3.pdf

Temkin, B. (2019). *The State of Customer ExperienceManagement, 2019. Benchmark of Customer Experience Activities, Competencies, and Maturity Levels*. Retrieved December 18, 2020, from https://www.qualtrics.com/docs/xmi/XMI_StateOfCustomerExperienceManagement-2019.pdf

Terpstra, T. (2011). Emotions, trust, and perceived risk: Affective and cognitive routes to floofd preparedness. *Risk Analysis*, *31*(10), 1658–1675. doi:10.1111/j.1539-6924.2011.01616.x PMID:21477090

Thales. (2018, January 8). KFC Use Facial Recognition for Payment in China. *Thales*. https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/magazine/kfc-use-facial-recognition-payment-china

Thalheim, B. (2013). *Entity-relationship modeling: Foundations of database technology*. Springer Science & Business Media.

The Economist Intelligence Unit. (2020). *The Internet of Things: Applications for Business. Exploring the transformative potential of IoT*. Author.

The Open Group. (2011). *Open Group Standard-TOGAF® Guide, Version 9.1*. The Open Group.

The Open Group. (2013b). *ArchiMate®.* Accessed and reviewed in April 2021, http://www.opengroup.org/subjectareas/*Entity*/archimate

The Open Group. (2014a). *The Open Group's Architecture Framework-Building blocks-Module 13*. Accessed and reviewed in April 2021, www.open-group.com/TOGAF

The Open Group. (2021). *The Open Group Professional Certification for the Data Scientist Profession (Open CDS).* The Open Group. Accessed and reviewed in April 2021, https://www.opengroup.org/certifications/certified-data-scientist-open-cds

Thomas & Thomas. (2018). Chatbot Using Gated End-to-End Memory Networks. *International Research Journal of Engineering and Technology, 5.*

Thomas, F., Handdon, L., Gilligan, R., Heinzmann, P., & De Gournay, C. (2005). Cultural Factors Shaping the Experience of ICTs: An Exploratory Review. *ResearchGate*. Obtenido de https://www.researchgate.net/publication/251783055_Cultural_Factors_Shaping_the_Experience_of_ICTs_An_Exploratory_Review

Thomsen, C. W. (1982). Robot Ethics and Robot Parody: Remarks on Isaac Asimov's *I, Robot* and Some Critical Essays and Short Stories by Stanislaw Lem. In T. P. Dunn & R. D. Erlich (Eds.), *The Mechanical God: Machines in Science Fiction* (pp. 19–26). Greenwood.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, *59*(2), 204–217. doi:10.5840/monist197659224 PMID:11662247

Tirunillai, S., & Tellis, G. J. (2014). Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *JMR, Journal of Marketing Research*, *51*(4), 463–479. doi:10.1509/jmr.12.0106

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Titz, K., Lanza-Abbott, J. A., & Cruz, G. C. Y. (2004). The anatomy of restaurant reviews: An exploratory study. *International Journal of Hospitality & Tourism Administration*, *5*(1), 49–65. doi:10.1300/J149v05n01_03

Tokman, M., Davis, L. M., & Lemon, K. N. (2007). The WOW factor: Creating value through win-back offers to reacquire lost customers. *Journal of Retailing*, *83*(1), 47-64. doi:10.1016/j.jretai.2006.10.005

To, M. L., & Ngai, E. W. (2006). Predicting the organisational adoption of B2C e-commerce: An empirical study. *Industrial Management & Data Systems*, *106*(8), 1133–1147. doi:10.1108/02635570610710791

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). *Processes of Technological Innovation*. Lexington books.

Torrent-Sellens, J. (2020). Industria 4.0 y resultados empresariales en España: un primer escaneado. *Oikonomics*. Retrieved December 18, 2020, from http://comein.uoc.edu/divulgacio/oikonomics/es/numero12/dossier/jtorrent.html

Torrent-Sellens, J. (2016). La economía del conocimiento y el conocimiento de la economía. *Oikonomics Revista de Economía. Empresa y Sociedad*, *5*, 26–32.

Trad, A., & Kalpić, D. (2020a). *Using Applied Mathematical Models for Business Transformation. Author Book*. IGI-Global. doi:10.4018/978-1-7998-1009-4

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, *14*(3), 207–222. doi:10.1111/1467-8551.00375

Trent, G. (2012). *The Rebirth of Mankind Homo Evolutis*. PDXdzyn.

Triantafillidou, A., & Siomkos, G. (2014). Consumption experience outcomes: Satisfaction, nostalgia intensity, word of-mouth communication and behavioural intentions. *Journal of Consumer Marketing*, *31*(6/7), 526–540. https://doi.org/10.1108/JCM-05-2014-0982

Trujillo, J., & Luj'an-Mora, S. (2003). *A UML Based Approach for Modelling ETL Processes in Data Warehouses*. Dept. de Lenguajes y Sistemas Inform'aticos. Universidad de Alicante.

Trunk, A., Birkel, H., & Hartmann, E. (2020). On the current state of combining human and artificial intelligence for strategic organizational decision making. *Business Research*, 1-45.

Tucker, C. (2018). Privacy, algorithms, and artificial intelligence. In *The Economics of Artificial Intelligence: An Agenda* (pp. 423–438). University of Chicago Press.

Tu, M. (2018). An exploratory study of Internet of Things (IoT) adoption intention in logistics and supply chain management. *International Journal of Logistics Management*, *29*(1), 131–151. doi:10.1108/IJLM-11-2016-0274

Turan, I., Şimşek, Ü., & Aslan, H. (2015). Eğitim araştırmalarında likert ölçeği ve likert tipi soruların kullanımı ve analizi. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, (30), 186–203.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *49*(236), 433–460. doi:10.1093/mind/LIX.236.433

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *49*, 433–460.

Turkish Renewable Energy Resources Support Mechanism (YEKDEM). (2011). *Official Gazette of the Republic of Turkey, 28782*. Retrieved from http: https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=18907&MevzuatTur=7&MevzuatTertip=5

Turletti, P. (2018). *El ROI de marketing y ventas, Cálculo y utilidad nuevo estandar de rendimiento*. ESIC.

Turner Lee, N. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information. Communication and Ethics in Society*, *16*(3), 252–260. doi:10.1108/JICES-06-2018-0056

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tušar, T., Gantar, K., Koblar, V., Ženko, B., & Filipič, B. (2017). A study of overfitting in optimization of a manufacturing quality control procedure. *Applied Soft Computing*, *59*, 77–87. doi:10.1016/j.asoc.2017.05.027

Tussyadiah, I. (2020). A review of research into automation in tourism: Launching the Annals of Tourism Research Curated Collection on Artificial Intelligence and Robotics in Tourism. *Annals of Tourism Research*, *81*, 102883. doi:10.1016/j.annals.2020.102883

Tynan, C., McKechnie, S., & Chhuon, C. (2010). Co-creating value for luxury brands. *Journal of Business Research*, *63*(11), 1156–1163. doi:10.1016/j.jbusres.2009.10.012

Uleman J. S. & Bargh J. A. (1989). *Unintended thought*. New York: Guilford Press.

Ulker-Demirel, E. (2019). The Features of New Communication Channels and Digital Marketing. In *Handbook of Research on Narrative Advertising* (pp. 302–313). IGI Global. doi:10.4018/978-1-5225-9790-2.ch026

Ullah, S., Akhtar, P., & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*, *71*, 69–78. doi:10.1016/j.indmarman.2017.11.010

UNCTAD. (2019). *Digitalization in Maritime Transport: Ensuring Opportunities for Development | UNCTAD*. United Nations Publication.

UNCTAD. (2020). Review of Maritime Transport 2020. *United Nations Publications*.

UNCTAD. (2021a). *Maritime profile: Brazil*. Retrieved from https://unctadstat.unctad.org/CountryProfile/MaritimeProfile/en-GB/076/index.html

UNCTAD. (2021b). Maritime profile: Portugal. *UNCTADstat*. Retrieved from https://unctadstat.unctad.org/CountryProfile/MaritimeProfile/en-GB/620/index.html

UNDESA. (2001). *Benchmarking e-government: A global perspective - assessing the progress of the un member states*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/English.pdf

UNDESA. (2003). *Un global e-government survey 2003*. Retrieved from https://publicadministration.un.org/egovkb/portals/egovkb/Documents/un/2003-Survey/unpan016066.pdf

UNDESA. (2004). *Un global e-government readiness report 2004: Towards access for opportunity*. Retrieved from https://publicadministration.un.org/egovkb/portals/egovkb/Documents/un/2004-Survey/Complete-Survey.pdf

UNDESA. (2005). *Un global e-government readiness report 2005: From e-government to e-inclusion* Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2005-Survey/Complete-survey.pdf

UNDESA. (2008). *United nations e-government survey 2008: From e-government to connected governance* Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2008-Survey/Complete-survey.pdf

UNDESA. (2010). *United nations e-government survey 2010: Leveraging e-government at a time of financial and economic crisis*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2010-Survey/Complete-survey.pdf

UNDESA. (2018). *United nations e-government survey 2018: Gearing e-government to support transformation towards sustainable and resilient societies*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2018-Survey/E-Government%20Survey%202018_FINAL%20for%20web.pdf

UNDESA. (2020). *E-government survey 2020: Digital government in the decade of action for sustainable development*. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2020-Survey/2020%20UN%20E-Government%20Survey%20(Full%20Report).pdf

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

United Nations. (2019). *Sustainable Development Goals*. Retrieved from https://sdgs.un.org/goals

Upchurch, M. (2018). Robots and ai at work: The prospects for singularity. *New Technology, Work and Employment*, *33*(3), 205–218. doi:10.1111/ntwe.12124

Upstart. (n.d.). https://www.upstart.com/about#

Urbinati, A., Bogers, M., Chiesa, V., & Frattini, F. (2019). Creating and capturing value from Big Data: A multiple-case study analysis of provider companies. *Technovation*, *84*, 21–36. doi:10.1016/j.technovation.2018.07.004

Urbinati, A., Chiaroni, D., Chiesa, V., & Frattini, F. (2020). The role of digital technologies in open innovation processes: An exploratory multiple case study analysis. *R & D Management*, *50*(1), 136–160. doi:10.1111/radm.12313

Urruticoechea, A., & Vernaza, E. (2018). *Una aplicación de la metodología biplot logísico: análisis de la sostenibilidad empresarial*. XI Semana Internacional de la Estadística y la Probabilidad. Facultad de Ciencias Físico Matemáticas Puebla.

US Congress. (2019). *HR 2231 – Algorithmic Accountability Act of 2019*. https://www.congress.gov/bill/116th-congress/house-bill/2231/text

Valohai, SigOpt, & Tecton. (2020). *Practical MLOps: How to get ready for production models*. Author.

Van Eck, N. J., & Waltman, L. (2011). *Text mining and visualisation using VOSviewer*. https://arxiv.org/pdf/1109.2058

van Geenhuizen, M. (2018). A framework for the evaluation of living labs as boundary spanners in innovation. *Environment and Planning C. Politics and Space*, *36*(7), 1280–1298. doi:10.1177/2399654417753623

Vanelslander, T., & Sys, C. (2020). *Maritime Supply Chains*. Elsevier Science.

Vargha, Z. (2018). Performing a strategy's world: How redesigning customers made relationship banking possible. *Long Range Planning*, *51*(3), 480–494.

Vargo, S. L., & Lusch, R. F. (2014). Inversions of service-dominant logic. *Marketing Theory*, *14*(3), 239–248.

Varian, H. R. (2014). *Intermediate Microeconomics A Modern Approach* (9th ed.). Theresia Kowara.

Vásquez, C., & Chik, A. (2015). "I am not a foodie…": Culinary capital in online reviews of Michelin restaurants. *Food & Foodways*, *23*(4), 231–250. doi:10.1080/07409710.2015.1102483

Vázquez Brust. (2020). *Ciencia de Datos para Gente Sociable. Una introducción a la exploración, análisis y visualización de datos.* Retrieved December 18, 2020, from https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/index.html

Vecchiato, G., Toppi, J., Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Bez, F., & Babiloni, F. (2011). Spectral EEG frontal asymmetries correlate with the experienced pleasantness of video commercial advertisements. *Medical & Biological Engineering & Computing*, *49*(5), 579–583. doi:10.100711517-011-0747-x PMID:21327841

Vedan, A. B., & Thomas, E. M. B. (2019). A Cabotagem Brasileira no Contexto Mundial: Uma Análise Compreensiva Da Lei N° 9.432/1997 e de seus desdobramento. In E. M. O. Martins & P. H. R. de Oliveira (Eds.), *Direito marítimo, portuário e aduaneiro: temas contemporâneos* (pp. 13–31). Arraes Editores.

Vélez-Rolón, A. M., Méndez-Pinzón, M., & Acevedo, O. L. (2020). Open Innovation Community for University-Industry Knowledge Transfer: A Colombian Case. *Journal of Open Innovation*, *68*(4), 1–17. doi:10.3390/joitmc6040181

Vella, A., Corne, D., & Murphy, C. (2009). *Hyper-heuristic decision tree induction*. Sch. of MACS, Heriot-Watt Univ., Edinburgh, UK. Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, *47*(1), 65–72.

Veneberg, R. (2014). *Combining enterprise architecture and operational data-to better support decision-making. University of Twente*. School of Management and Governance.

Venkatesh, V. G., Zhang, A., Deakins, E., Mani, V., & Shi, Y. (2020). Supply Chain Integration Barriers to Port-Centric Logistics—An Emerging Economy Perspective. *Transportation Journal*, *59*(3), 215–253. doi:10.5325/transportationj.59.3.0215

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *Management Information Systems Quarterly*, *36*(1), 157–178. doi:10.2307/41410412

Verbeek, P. P. (2006). Materializing morality: Design ethics and technological mediation. *Science, Technology & Human Values*, *31*(3), 361–380. doi:10.1177/0162243905285847

Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press. doi:10.7208/chicago/9780226852904.001.0001

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. *Journal of Retailing*, *85*(1), 31–41.

Verkijika, S. F., & De Wet, L. (2018). A usability assessment of e-government websites in sub-saharan africa. *International Journal of Information Management*, *39*, 20–29. doi:10.1016/j.ijinfomgt.2017.11.003

Verwer, S., & Calders, T. (2013). Introducing positive discrimination in predictive models. In *Discrimination and privacy in the information society* (pp. 255–270). Springer. doi:10.1007/978-3-642-30487-3_14

Vetrò, A., Santangelo, A., Beretta, E., & De Martin, J. C. (2019). AI: From rational agents to socially responsible agents. *Digital Policy*. *Regulation & Governance*, *21*(3), 291–304. Advance online publication. doi:10.1108/DPRG-08-2018-0049

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, *28*(2), 118–144.

Viana, J., & Peralta, H. (2021). Online Learning: From the Curriculum for All to the Curriculum for Each Individual. *Journal of New Approaches in Educational Research*, *10*(1), 122–136. doi:10.7821/naer.2021.1.579

Vicente Cuervo, M. R., & López Menéndez, A. J. (2006). A multivariate framework for the analysis of the digital divide: Evidence for the european union-15. *Information & Management*, *43*(6), 756–766. doi:10.1016/j.im.2006.05.001

Vicente Galindo, P., Vaz, E., & de Noronha, T. (2015). How Corporations Deal with Reporting Sustainability: Assessment Using the Multicriteria Logistic Biplot Approach. *Systems*, *3*(1), 6–26. doi:10.3390ystems3010006

Vicente Villardón, J.L. (2015). *Multbiplot: A package for Multivariate Analysis using Biplots. Departamento de Estadística.* Universidad de Salamanca. http://biplot.usal.es/ClassicalBiplot/index.html

Vicente-Villardon, J. L., Galindo-Villardon, M. P., & Blazquez-Zaballos, A. (2006). *Logistic Biplots in Multiple Correspondence Analysis and Related Methods*. Statistics in Social and Behavioural Sciences Series.

Vickers, A. J. (2005). Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. *BMC Medical Research Methodology*, *5*(1), 35. doi:10.1186/1471-2288-5-35 PMID:16269081

Vieira, J., & Fialho, G. (2020). Port Management Modernization and Integrated Operational Planning. *Revista Eletrônica de Estratégia & Negócios*, *13*, 196. doi:10.19177/reen.v13e0I2020196-224

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vigderhous, G. (1977). The level of measurement and "permissible" statistical analysis in social research. *Pacific Sociological Review*, *20*(1), 61–72. doi:10.2307/1388904

Vilner, Y. (n.d.). *Chatbots101:The Evolution of Customer Retention Latest Trend*. Available: https://www.entrepreneur.com/article/293439

Vincent, J. (2017, January 4). The UK's national health service is testing out a medical chatbot as a non-emergency helpline. *The Verge*. https://www.theverge.com/2017/1/4/14168312/uk-nhs-babylon-medical-chatbot-helpline

Visengeriyeva, L., Kammer, A., Bär, I., Kniesz, A., & Plöd, M. (n.d.). *Machine learning operations*. Retrieved from https://ml-ops.org/

Vogelsang, A., & Borg, M. (2019). *Requirements engineering for machine learning: Perspectives from data scientists.* Paper presented at the IEEE 27th International Requirements Engineering Conference Workshops (REW). 10.1109/REW.2019.00050

Voice2Biz. (n.d.). *Google Home and Alexa Talk to aach other & have a conversation*. https://www.voice2biz.com/google-home-and-alexa-talk-to-each-other-have-a-conversation/

Vrontis, D., Thrassou, A., & Czinkota, M. R. (2011). Wine marketing: A framework for consumer-centred planning. *Journal of Brand Management*, *18*(4-5), 245–263. doi:10.1057/bm.2010.39

Vroom, V. H. (1964). *Work and motivation*. Wiley.

Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.

Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox: Implications for governance and communication of natural hazards. *Risk Analysis*, *33*(6), 1049–1065. doi:10.1111/j.1539-6924.2012.01942.x PMID:23278120

Wadgave, U., & Khairnar, M. R. (2016). Parametric tests for Likert scale: For and against. *Asian Journal of Psychiatry*, *24*, 67–68. doi:10.1016/j.ajp.2016.08.016 PMID:27931911

Wahlstrom, K., Ul-haq, A., & Burmeister, O. (2020). Privacy by design. *AJIS. Australasian Journal of Information Systems*, *24*, 1–9. doi:10.3127/ajis.v24i0.2801

Wallach, W., Franklin, S., & Allen, C. (2010). A Conceptual and Computational Model of Moral Decision Making in Human and Artificial Agents. *Trends in Cognitive Sciences*, *2*(3), 454–485. PMID:25163872

Waller, M. A., & Fawcett, S. E. (2013). Click here for a data scientist: Big data, predictive analytics, and theory development in the era of a maker movement supply chain. *Journal of Business Logistics*, *34*(4), 249–252. doi:10.1111/jbl.12024

Walls, A. R., Okumus, F., Wang, Y., & Kwun, D. J. W. (2011). An epistemological view of consumer experience. *International Journal of Hospitality Management*, *30*(1), 10–21.

Wang, B., Yang, Z., Xuan, J., & Jiao, K. (2020). Crises and opportunities in terms of energy and AI technologies during the COVID-19 pandemic. *Energy AI*, *1*, 100013. doi:10.1016/j.egyai.2020.100013

Wang, C. H., & Lien, C. Y. (2019). Combining design science with data analytics to forecast user intention to adopt customer relationship management systems. *Journal of Industrial and Production Engineering*, *36*(4), 193–204. doi:10.1080/21681015.2019.1645051

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wang, D., Weisz, J. D., Muller, M., Ram, P., Geyer, W., Dugan, C., Tausczik, Y., Samulowitz, H., & Gray, A. (2019). Human-AI collaboration in data science: Exploring data scientists' perceptions of automated AI. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1-24. 10.1145/3359313

Wang, J., & Karypis, G. (2004). Summary: Efficiently Summarizing Transactions for Clustering. In *Proceedings of the Fourth IEEE International Conference on Data Mining*. IEEE Computer Society. 10.1109/ICDM.2004.10105

Wang, K., Xu, C., & Liu, B. (1999). Clustering Transactions Using Large Items. In *Proceedings of the Eighth International Conference on Information and Knowledge*. Association for Computing Machinery. 10.1145/319950.320054

Wang, L., & Alexander, C. A. (2016). Machine learning in big data. International *Journal of Mathematical*. *Engineering and Management Sciences*, *1*(2), 52–61.

Wang, L., Luo, G. L., Sari, A., & Shao, X. F. (2020). What nurtures fourth industrial revolution? An investigation of economic and social determinants of technological innovation in advanced economies. *Technological Forecasting and Social Change*, *161*, 120305. doi:10.1016/j.techfore.2020.120305

Wang, P., Guo, J., Lan, Y., & Cheng, X. (2014). Modeling Retail Transaction Data for Personalized Shopping Recommendation. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*. Association for Computing Machinery. 10.1145/2661829.2662020

Wang, Q., Li, S., & Pisarenko, Z. (2020). Heterogeneous effects of energy efficiency, oil price, environmental pressure, R&D investment, and policy on renewable energy – evidence from the G20 countries. *Energy*, *209*, 118322. doi:10.1016/j.energy.2020.118322

Wang, Q., Valchuis, L., Thompson, E., Conner, D., & Parsons, R. (2019). Consumer Support and Willingness to Pay for Electricity from Solar, Wind, and Cow Manure in the United States: Evidence from a Survey in Vermont. *Energies*, *12*(23), 4467. doi:10.3390/en12234467

Wang, R., Luo, J., & Huang, S. S. (2020). Developing an artificial intelligence framework for online destination image photos identification. *Journal of Destination Marketing & Management*, *18*, 100512. doi:10.1016/j.jdmm.2020.100512

Wang, S.-C., Soesilo, P. K., & Zhang, D. (2015). Impact of luxury brand retailer co-branding strategy on potential customers: A cross-cultural study. *Journal of International Consumer Marketing*, *27*(3), 237–252. doi:10.1080/08961530.2014.970320

Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems with Applications*, *143*, 113042. doi:10.1016/j.eswa.2019.113042

Wang, X., Nguyen, K., & Nguyen, B. P. (2020). Churn prediction using ensemble learning. In *ACM International Conference Proceeding Series* (pp. 56–60). 10.1145/3380688.3380710

Ward, R., & Bjørn-Andersen, N. (2021). The Origins of Maritime Informatics. In Maritime Informatics. Springer. doi:10.1007/978-3-030-50892-0_1

Warwick, K. (2002). *I, Cyborg*. Century.

Washizaki, H., Uchida, H., Khomh, F., & Guéhéneuc, Y.-G. (2020). *Machine learning architecture and design patterns*. Academic Press.

Wason, H., & Charlton, N. (2015). How positioning strategies affect co-branding outcomes. *Cogent Business & Management*, *2*(1), 1092192. doi:10.1080/23311975.2015.1092192

WCO. (2021). *Managing Transition to a Single Window, Volume 1, Part IX*. World Customs Organization.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Webb, H., Patel, M., Rovatsos, M., Davoust, A., Ceppi, S., Koene, A., Dowthwaite, L., Portill, V., Jirotka, M., & Cano, M. (2019). It would be pretty immoral to choose a random algorithm: Opening up algorithmic interpretability and transparency. *Journal of Information Communication and Ethics in Society*, *17*(2), 210–228. doi:10.1108/JICES-11-2018-0092

Weber, R. H. (2010). Internet of Things–New security and privacy challenges. *Computer Law & Security Review*, *26*(1), 23–30. doi:10.1016/j.clsr.2009.11.008

Weem, G. H., Onwuegbuzie, A. J., & Collins, K. M. T. (2006). The role of reading comprehension in responses to positively and negatively worded items on rating scales. *Evaluation and Research in Education*, *19*(1), 3–20. doi:10.1080/09500790608668322

Weems, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement & Evaluation in Counseling & Development*, *34*(3), 166–176. doi:10.1080/07481756.2002.12069033

Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). A review of machine learning for the optimization of production processes. *International Journal of Advanced Manufacturing Technology*, *104*(5-8), 1889–1902. doi:10.100700170-019-03988-5

Weinberg, B. D., Milne, G. R., Andonova, Y. G., & Hajjat, F. M. (2015). Internet of Things: Convenience vs. privacy and secrecy. *Business Horizons*, *58*(6), 615–624. doi:10.1016/j.bushor.2015.06.005

Weinmann, M., Schneider, C., & Vom Brocke, J. (2016). Digital Nudging. *Business & Information Systems Engineering*, *433-436*. Advance online publication. doi:10.2139srn.2708250

Weld, D. S., & Bansal, G. (2019). The challenge of crafting intelligible intelligence. *Communications of the ACM*, *62*(6), 70–79. doi:10.1145/3282486

Weng, L., & Cheng, C. (2000). Effects of response order on Likert type scales. *Educational and Psychological Measurement*, *60*(6), 908–924. doi:10.1177/00131640021970989

Werbos, P. J. (1982). Applications of advances in nonlinear sensitivity analysis. In R. F. Drenick & F. Kozin (Eds.), *System Modeling and Optimization* (pp. 762–770). Springer Berlin Heidelberg. doi:10.1007/BFb0006203

West, D. M. (2003). *Global e-government, 2003*. Retrieved from http://www.insidepolitics.org/egovt03int.pdf

West, D. M. (2004b). *Global e-government, 2004*. Retrieved from http://www.insidepolitics.org/egovt04int.pdf

West, D. M. (2005). *Global e-government, 2005*. Retrieved from http://www.insidepolitics.org/egovt05int.pdf

West, D. M. (2006). *Global e-government, 2006*. Retrieved from http://www.insidepolitics.org/egovt06int.pdf

West, D. M. (2007). *Global e-government, 2007*. Retrieved from http://www.insidepolitics.org/egovt07int.pdf

Westbrook, R. A., & Oliver, R. L. (1991). The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction. *The Journal of Consumer Research*, *18*(1), 84–91. doi:10.1086/209243

West, D. M. (2004a). E-government and the transformation of service delivery and citizen attitudes. *Public Administration Review*, *64*(1), 15–27. doi:10.1111/j.1540-6210.2004.00343.x

Weston, Bordes, Rush, van Merrienboer, Joulin, & Mikolov. (n.d.). Towards Ai-complete Question Answering: A Set Of Prerequisite Toy Tasks. In *Soft Computing for Problem Solving Arch.* Academic Press.

Weston, J., Bordes, A., & Chopra, S. (2015). *Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks*. arXiv preprint, arXiv: 1502.05968

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Weston, J., Chopra, S., & Bordes, A. (2015). Memory networks. *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.

Whitmore, A. (2012). A statistical analysis of the construction of the united nations e-government development index. *Government Information Quarterly*, *29*(1), 68–75. doi:10.1016/j.giq.2011.06.003

Wieckowski, A. (2019). When neuromarketing crosses the line. *Harvard Business Review*. Accessed from: https://hbr.org/2019/01/when-neuromarketing-crosses-the-line?ab=at_articlepage_recommendedarticles_bottom1x1

Wierenga, B. (2010). Marketing and artificial intelligence: Great opportunities, reluctant partners. In *Marketing intelligent systems using soft computing* (pp. 1–8). Springer.

Wiggers, K. (2019, March 19). Ojo Labs raises $45 million to develop a chatbot for real estate. *VentureBeat*. https://venturebeat.com/2019/03/19/ojo-labs-raises-45-million-to-develop-a-chatbot-for-real-estate/

Wilber, K. (1985). *The Holographic Paradigm and other paradox, exploring the leading edge of science*. New Science Library.

Wilson, H. J., Daugherty, P., & Bianzino, N. (2017). The jobs that artificial intelligence will create. *MIT Sloan Management Review*, *58*(4), 14.

Winston, P. H. (1992). *Artificial Intelligence* (3rd ed.). Addison-Wesley.

Wirth, N. (2018). Hello marketing, what can artificial intelligence help you with? *International Journal of Market Research*, *60*(5), 435–438. doi:10.1177/1470785318776841

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, *29*(5), 907–931. doi:10.1108/JOSM-04-2018-0119

Witten, I. H., & Frank, E. (2002). Data mining: Practical machine learning tools and techniques with java implementations. *SIGMOD Record*, *31*(1), 76–77. doi:10.1145/507338.507355

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factory analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 189–194. doi:10.100710862-005-9004-7

World Bank. (2020). *World bank country and lending groups*. Retrieved from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Economic Forum. (2019). *What are ICTs?* Retrieved from en.weforum.org: https://es.weforum.org/agenda/2019/02/que-son-las-tics/

World Shipping Council. (2021). *Industry Issues*. Retrieved from https://www.worldshipping.org/industry-issues

Wu. (2020). *Bulk ship routing and scheduling under uncertainty* (Doctoral dissertation). Hong Kong Polytechnic University, China. Retrieved from https://theses.lib.polyu.edu.hk/handle/200/10660

Wu, D., Jennings, C., Terpenny, J., Gao, R., & Kumara, S. (2017). A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. *Journal of Manufacturing Science and Engineering*, *139*(7), 071018. Advance online publication. doi:10.1115/1.4036350

Xiao, W., Zhao, H., Pan, H., Song, Y., Zheng, V. W., & Yang, Q. (2019, July). Beyond personalization: Social content recommendation for creator equality and consumer satisfaction. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 235-245). ACM.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Xiong, C., Merityand, S., & Socher, R. (2016). *Dynamic Memory Networks for Visual and Textual Question Answering*. ICML.

Xue, S., Wu, M., Kolen, J., Aghdaie, N., & Zaman, K. A. (2017). Dynamic Difficulty Adjustment for Maximized Engagement in Digital Games. *WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web*.

Xu, H., Luo, X., Carroll, J. M., & Rosson, M. B. (2011). The Personalization Privacy Paradox. *Decision Support Systems*, *51*(1), 42–52. doi:10.1016/j.dss.2010.11.017

Yadav, M., Perumal, M., & Srinivas, M. (2020). Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos, Solitons, and Fractals*, *139*, 110050. doi:10.1016/j.chaos.2020.110050 PMID:32834604

Yams, N. B., Richardson, V., Shubina, G. E., Albrecht, S., & Gillblad, D. (2020). Integrated AI and Innovation Management: The Beginning of a Beautiful Friendship. *Technology Innovation Management Review, 10*(11).

Yan, E. (2019). *Data science and agile: What works and what doesn't work*. Retrieved from https://towardsdatascience.com/agiledatascience-3b7ca65278a4

Yang, C., Dillard, J., & Li, R. (2018). Understanding fear of Zika: Personal, interpersonal, and media Influences. *Risk Analysis*, *38*(12), 2535–2545. doi:10.1111/risa.12973 PMID:29392760

Yang, S. B., Hlee, S., Lee, J., & Koo, C. (2017). An empirical examination of online restaurant reviews on Yelp. com. *International Journal of Contemporary Hospitality Management*, *29*(2), 817–839. doi:10.1108/IJCHM-11-2015-0643

Yang, S., & Park, S. (2020). The effects of renewable energy financial incentive policy and democratic governance on renewable energy aid effectiveness. *Energy Policy*, *145*, 111682. doi:10.1016/j.enpol.2020.111682

Yan, H., Chen, K., & Liu, L. (2006). Efficiently Clustering Transactional Data with Weighted Coverage Density. In *Proceedings of the 15th International Conference on Information and Knowledge Management*. Association for Computing Machinery. 10.1145/1183614.1183668

Yeşilkanat, C. M. (2020). Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine-learning algorithm. *Chaos, Solitons, and Fractals*, *140*, 110210. doi:10.1016/j.chaos.2020.110210 PMID:32843823

Yoo, C. W. (2020). An Exploration of the Role of Service Recovery in Negative Electronic Word-of-Mouth Management. *Information Systems Frontiers*, *22*(3), 719–734. doi:10.100710796-018-9880-5

Yoon, S., Feng, J., & Jain, A. K. (2010). On Latent Fingerprint Enhancement. *Proc. SPIE Biometric Technology for Human Identification VII*.

Yu, H., Shen, Z., & Miao, C. (2018). Building Ethics into Artificial Intelligence, *Computer Science, Artificial Intelligence*. Cornell University. https://arxiv.org/abs/1812.02953

Yue, B. (2020). Topological Data Analysis of Two Cases: Text Classification and Business Customer Relationship Management. *Journal of Physics: Conference Series*, *1550*(3), 032081. Advance online publication. doi:10.1088/1742-6596/1550/3/032081

Yu, F., Liu, Q., Wu, S., Wang, L., & Tan, T. (2016). A Dynamic Recurrent Model for Next Basket Recommendation. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery. 10.1145/2911451.2914683

Zahedi, M., & Ghadi, O. R. (2015). Combining Gabor filter and FFT for fingerprint enhancement based on a regional adaption method and automatic segmentation. *Signal, Image and Video Processing*, *9*(2), 267–275. doi:10.100711760-013-0436-3

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. doi:10.1037/h0025848

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *The American Psychologist*, *35*(2), 151–175. doi:10.1037/0003-066X.35.2.151

Zarsky, T. Z. (2016). Incompatible: The GDPR in the age of big data. *Seton Hall Law Review*, *47*, 995.

Zhang, D., Zhang, Y., Li, Q., Plummer, T., & Wang, D. (2019, July). Crowdlearn: A crowd-ai hybrid system for deep learning-based damage assessment applications. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)* (pp. 1221-1232). IEEE. 10.1109/ICDCS.2019.00123

Zhang, J. R. L.-C. (2013). Adaptive Directional Total-Variation Model for. IEEE Transactions on Information Forensics And Security, 8.

Zhang, T., Gao, C., Ma, L., Lyu, M., & Kim, M. (2019). *An empirical study of common challenges in developing deep learning applications.* Paper presented at the 2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE). 10.1109/ISSRE.2019.00020

Zhang, C. B., & Li, Y. N. (2019). How social media usage influences B2B customer loyalty: Roles of trust and purchase risk. *Journal of Business and Industrial Marketing*.

Zhang, J., Harman, M., Ma, L., & Liu, Y. (2020). *Machine learning testing: Survey, landscapes and horizons. IEEE Transactions on Software Engineering*. Early Access.

Zhang, J., Lai, R., & Kuo, C. C. J. (2013). Adaptive directional total-variation model for latent fingerprint segmentation. *IEEE Transactions on Information Forensics and Security*, *8*(8), 1261–1273. doi:10.1109/TIFS.2013.2267491

Zhang, X., & Savalei, V. (2016). Improving the factor structure of psychological scales: The Expanded format as an alternative to the Likert scale format. *Educational and Psychological Measurement*, *76*(3), 357–386. doi:10.1177/0013164415596421 PMID:27182074

Zhang, Z., Zhang, Z., & Law, R. (2014). Relative importance and combined effects of attributes on customer satisfaction. *Service Industries Journal*, *34*(6), 550–566. doi:10.1080/02642069.2014.871537

Zhao, P., Lu, Z., Fang, J., Paramati, S. R., & Jiang, K. (2020). Determinants of renewable and non-renewable energy demand in China. *Structural Change and Economic Dynamics*, *54*, 202–209. doi:10.1016/j.strueco.2020.05.002

Zhao, Y., Tang, K.-K., & Wang, L.-L. (2013). Do renewable electricity policies promote renewable electricity generation? Evidence from panel data. *Energy Policy*, *62*, 887–897. doi:10.1016/j.enpol.2013.07.072

Zhao, Y., Xu, X., & Wang, M. (2019). Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews. *International Journal of Hospitality Management*, *76*, 111–121. doi:10.1016/j.ijhm.2018.03.017

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350–361. doi:10.1016/j.neucom.2017.01.026

Zickfeld, J. H., Schubert, T. W., Herting, A. K., Grahe, J., & Faasse, K. (2020). Correlates of health-protective behavior during the initial days of the COVID-19 outbreak in Norway. *Frontiers in Psychology*, *11*, 11. doi:10.3389/fpsyg.2020.564083 PMID:33123045

Zinkevich, M. (2017). *Rules of machine learning: Best practices for ML engineering*. Retrieved from https://developers.google.com/machine-learning/guides/rules-of-ml

Žliobaitė, I. (2015). *A survey on measuring indirect discrimination in machine learning.* arXiv preprint arXiv:1511.00148.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zomerdijk, L. G., & Voss, C. A. (2011). NSD processes and practices in experiential services. *Journal of Product Innovation Management*, *28*(1), 63–80.

Zwanka, R., & Buff, C. (2020). Generation COVID-19: A conceptual framework of consumer behavioral changes to be caused by the COVID-19 pandemic. *Journal of International Consumer Marketing*, *33*(1), 58–67. doi:10.1080/08961530.2020.1771646

Żyminkowska, K. (2019). Concepts of Customer Activism. In *Customer Engagement in Theory and Practice* (pp. 1–22). Palgrave Pivot. doi:10.1007/978-3-030-11677-4_1

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

About the Contributors

Valentina Chkoniya received her PhD in Technical Sciences by Murmansk State Technical University, Russia, in 2003. Actually, she is a professor at the University of Aveiro, member of GOVCOPP, Portugal, and managing partner of VT Mar (company in the marketing intelligence market). Valentina Chkoniya is a highly professional and experienced professor with a deep understanding and knowledge regarding the business, based on a deep strategic perspective and building on more than 20 years of experience. Her current research interests include data analysis, decision support systems, marketing research, consumer behaviour and information technologies for retail.

* * *

Clony Abreu Junior is a research fellow. Graduated in Computer Science, at the Ruy Barbosa College, Salvador, Bahia – Brazil (2005), postgraduate in IT Management by Metropolitan College of Belo Horizonte – MG, Brazil (2008), MBA in Applications Project for Mobile Devices by Instituto Newton Paiva Ferreira de Belo Horizonte – MG (2017). Master in Computer Engineering at the Aveiro University, Portugal (2021).

David Aguilar Del Carpio is a professor at the Faculty of Economics of the Universidad Nacional de San Agustín de Arequipa. Economist graduated from the same university, Master in Economics from the PUCP, graduated from the graduate school of the UNSA of the Master in Formulation, Evaluation and Management of Investment Projects. Experience in research of public policies and social problems, as well as in positions of human resource management, planning and spending in the public and private sectors. Skill in indicator analysis, statistics and information consolidation.

Abdullah M. M. Altrad is an Associate Professor at the American University of Phnom Penh (AUPP). Received Master of IT in 2011 from University Utara Malaysia (UUM), and also obtained PhD of IT from School of Computing - UUM as well. Also, worked as research assistant in UUM (2014) for Broadband Service Project of Last-mile connection-based Powerline Communication Network, then moved to Almadinah International University Malaysia as a lecturer of Computer Science since 2015 until August 2020. Moreover, introduced an Adaptive Recourse Allocation Technique for Powerline Communication Network based on Combinatorial Optimization and Metaheuristic Techniques. The current research topics are on Machine Learning, Wireless Networks, IoT, and Cybersecurity.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Muhammad Ashfaq is a post-doctorate researcher at Shenzhen University, China. In 2020, he obtained his PhD in Marketing from Dongbei University of Finance and Economics, China. His work has been published in different SSCI-ranked journals. His research includes consumer behavior, technology adoption, and digital marketing.

Elodie Attié is currently a scientific project manager at Capgemini Engineering. She obtained a PhD in marketing about the acceptance of the Internet of Things and its influence on consumer well-being in 2019. Her research interests relate to innovations, marketing, consumer behavior and well-being.

Anteneh Ayanso is Professor of Information Systems and Director of the Centre for Business Analytics at Brock University. He teaches Business Analytics, Database Design and Management, and Data Mining Techniques & Applications. His research interests focus primarily on data management, business analytics, electronic commerce, and electronic government. He has published many articles in leading journals such as Decision Sciences, Decision Support Systems, European Journal of Operational Research, Journal of Database Management, International Journal of Electronic Commerce, Government Information Quarterly, among others. Dr. Ayanso's research has been funded by internal and external grants, including a Discovery Research Grant by the Natural Sciences and Engineering Research Council (NSERC) of Canada. He is currently serving as an Associate Editor at Decision Support Systems journal and a review board member at Journal of Database Management, International Journal of Convergence Computing, and International Journal of Electronic Commerce.

Veena Bansal is an Associate Professor with the Indian Institute of Technology Kanpur. She is widely published in the area of enterprise resource planning. Her current area of interest include Big Data, Machine Learning applications, Data Science Adoption.

Tamer Baran is PhD lecturer. Dr. Baran hold his bachelor degree from Dumlupinar University, master and PhD degrees in Marketing from Pamukkale University. His research field is marketing and specified areas are Islamic marketing, marketing research, consumer behavior, branding and retailing. Studies of Dr. Baran has published in prestigious academic journals in these areas. He has worked in Pamukkale University, Kale VS since 2011.

Belém Barbosa is an Assistant Professor at the School of Economics and Management, University of Porto. She is also a Researcher at the Research Unit of Governance, Competitiveness and Public Policy (GOVCOPP). Her research main research interests are digital marketing and consumer behavior.

Borja Barinaga is a digital art professor in the videogames department at the Francisco de Vitoria University in Madrid. He is an official researcher at the University Complutense in Madrid, and he is a member of the research group Museum I+D+C, Laboratory for the Digital Culture and Hypermedia Museography. Barinaga wrote the first book about video game design in the Spanish language and also is one of the founders of the first video game degree in Spain. His research field is digital languages and communication.

Mª Manuela Batista is a specialist in Transport Services, Postgraduate in Environmental Port Management, Coordinator of Port Management Undergraduate Course at ENIDH (Escola Superior

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Náutica Infante D. Henrique - Portuguese Maritime College). Professor at Transport and Logistics Department (Applied Physical Chemistry and Pollution – Master Course in Deck and Bridge Operations; Environmental Management in Transport - Degree in Port Management and Degree in Transport and Logistics Management; Applied Chemistry and Physical Chemistry - Degree in Deck and Bridge Operations). Participates in national / European projects in the maritime-port sector and coordinates a Training Academy (MSC-ENIDH Academy) dedicated to Maritime Transport. Since 2001 coordinates and teaches several training actions/courses: Dangerous and Noxious Cargo Operations; Familiarization / Specialization for the transportation of Goods in Tankers (Oil Tankers, Chemicals and Liquefied Gas). Is a consultant in companies / associations dedicated to the transport, handling and storage of chemical products / dangerous goods.

Liliana Bernardino has a Masters Degree in Mathematics & Statistics with a following Master Degree in Machine Learning and Decision Support Systems. With more than 16 years' experience of using simple and complex analytics to change data into knowledge by focusing her work and efforts in customer insights and strategic/behavioral analytics. She's passionate about solving questions like "so what?", "how it happened?" and "why?" in consumer data to create true customer insight. She has significant experience delivering and drawing advanced predictive analytics solutions to drive customer loyalty into commercial results. Has the belief that if a company applies analytics correctly this turns out not just into measurement tools, but into an insight finder, a strategy builder and a loyalty weapon. Throughout her career she has built (and still do) platforms and statistical solutions that were the "ramp up" for the customer knowledge growth and in the past few years leads CRM, Data Science and Report Building Teams in Sonae MC

Chintan Bhatt (Ph.D.) is currently working as an Assistant Professor in Computer Engineering department, Chandubhai S. Patel Institute of Technology, CHARUSAT. He is a member of IEEE, EAI, ACM, CSI, AIRCC and IAENG (International Association of Engineers). His areas of interest include Internet of Things, Data Mining, Web Mining, Networking, Security Mobile Computing, Big Data and Software Engineering. He has more than 5 years of teaching experience and research experience, having good teaching and research interests. He has chaired a track in CSNT 2015 and ICTCS 2014. He has been working as Reviewer in Wireless Communications, IEEE (Impact Factor-6.524) and Internet of Things Journal, IEEE, Knowledge-Based Systems, Elsevier (Impact Factor-2.9) Applied Computing and Informatics, Elsevier and Mobile Networks and Applications, Springer. He has delivered an expert talk on Internet of Things at Broadcast Engineering Society Doordarshan, Ahmedabad on 30/09/2015. He has been awarded Faculty with Maximum Publication in CSIC Award and Paper Presenter Award at International Conference in CSI-2015, held at New Delhi.

Hector Bienvenido is currently an assistant professor in the area of Social Research Methods at Complutense University of Madrid. Héctor's research interests focus on the social implications of ICTs (mobile phones, social networks) and videogames, especially in relation to the processes of interaction, social exclusion, agencies and emergency; as well as the study of youth and leisure practices and cultures, virtual communities, social change and user experiences (UX). He is a member of the research group 'Ordinary Sociology/Sociología Ordinaria' (UCM). Héctor has published numerous conference and journal papers such as papers in DiGRA, Journal of Critical Care, and Revista Española de Investigaciones Sociológicas.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jacques Bughin has lectured economics and management since 1992, and has published tens of research in outlets such as Management Science, Research Policy, IEEE, Eutropean Economic Review, Harvard Business Review and Sloan Management Review. After retiring as senior partner at McKinsey and as Director of the McKinsey Global Institute, he now advises various companies on strategies, and the UN on digitigitization, is Venture Partner at Antler, and Senior Advisor to Fortino Capital. He is also part of the Technology advisor board of the European Parliament.

Yezelia Cáceres is a professor at the Faculty of Economics of the Universidad Nacional de San Agustín de Arequipa, Peru. She recently published: 'Informal food systems and differential mobility during the COVID-19 pandemic in Arequipa, Peru together with Aaron Malone and Anabel Taya (https://doi.org/10.3828/tpr.2020.61). On the other hand, her book chapter stands out: Alpaca Meat: Constraints and Reinforcements on Consumption, in book: Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior (pp. 286-302) DOI: 10.4018 / 978-1-7998-3115- 0.ch016, prepared with Sara Gonzales and Juan Rivera. Among other contributions, there is the Analysis of the instrument for measuring institutional capacity. Case study: Pampacolca 2018 Available at the CIES Economic and Social Research Consortium. Theoretical contributions for the analysis of competitiveness and the value chain in micro and small companies. Ilustro 2018, vol 9 available at https://doi.org/10.36901/illustro.v9i0.1224 The use of information and communication technologies to improve agricultural production in the Arequipa region 2014 - Research Contest Research Consortium Económica y Social CIES ¨The use of ICT for key industries in the southern economic corridor: alpaca and coffee. Enhancing the Competitiveness of SMEs in the Southern Economic Corridor in Peru (2012) KDS Korean Development Strategy Using the radio to improve the local responses to climate variability, Center for Development and Informatics (CDI), University of Manchester, United Kingdom (2012).

José Luís Cacho is CEO of APS – Ports of Sines and the Algarve Authority, S.A. and ex-President and Adviser of APLOP - Association of Portuguese Speaking Ports. The Port of Sines is an open deep-water sea port with excellent maritime access, without restrictions, leading the Portuguese port sector in the volume of cargo handled, and offering unique natural characteristics to receive any type of vessels. He also was member of the Monitoring Committee of the Logistics Platforms National Network Plan, a member of the Monitoring Group for the rail link to the Port of Aveiro and of the Committee on Accreditation of Civil Engineering Graduation at the University of Trás-os -Montes e Alto Douro (UTAD), by the Engineers Order between. He was also member of the Biology Department Council, of the University of Aveiro, while being a member of the Council General of the University of Coimbra. He was CEO of APA - Port of Aveiro Authority, S.A., where he had already been Member of the Board of Directors and CEO of APFF - Port of Figueira da Foz Authority, S.A. and President of APP – Portugal Ports Association.

Michele Cincera is Professor of Industrial economics at the Solvay Brussels School of Economics and Management – Université Libre de Bruxelles. Since 2012, he is the Director of the International Centre of Innovation, Technology and Education Studies (iCite). His research interests and expertise embrace the quantitative assessment at the micro-level of innovative and entrepreneurial activities, their determinants and socio-economic impacts as well as the analysis of National Innovation Systems and the socio-economic impact of policies supporting science, research and innovation. He is currently member

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the scientific board of the Brussels Institute for Statistics and Analysis (IBSA) and member of the international jury of the pôles de compétitivité wallons.

Kimberly Rose Clark is a Senior Lecturer and Researcher at Dartmouth College. She also serves as Co-Founder and Chief Research Officer of Merchant Mechanics and Founding Partner of Bellwether Citizen Response. Dr. Clark is a successful "pracademic" focused on human decision making. Her research specialty centers on applied behavioral neuroscience research with 30+ years of designing and implementing research programs focused on science communications.

Luís Cruz is an Assistant Professor at the Delft University of Technology and the Scientific Manager of the AI-for-Fintech Research lab. He holds a Ph.D. in Computer Science from the University of Porto, Portugal. Dr. Cruz has been actively contributing to the research fields of Machine Learning Engineering and Sustainable Software Engineering. His work aims at defining software engineering practices for machine learning projects and at improving development processes to deliver carbon-efficient software.

Pedro Gusmão is a student and researcher at University of Aveiro.

Fernando Cruz Goncalves is a specialist in Administration and Management - Maritime-Port Economics (ENIDH / Escola Naval / ISCIA), Master in Maritime Management and Transport Economics (ITMMA - Antwerp University), MBA in Logistics Management (IST / ISCTE / EGP), Degree in Management and Technologies Maritime (ENIDH), Degree in Administration and Maritime Management (ENIDH), Bachelor in Pilotage (ENIDH). Merchant Marine Officer, Coordinator of the degree in Port Management, Coordinator of the degree in Transport and Logistics Management and Coordinator of the Master in Port Management at ENIDH. Professor at the Department of Transport and Logistics. Consultant in Transport, Ports and Logistics in several projects and studies carried out in the scope of business and programs co-financed by the European Commission. Author and co-author of several publications in the area of Management, Economics of Transport and Logistics. Author of the book the "Maritime Transport of Containerized Cargo Market" (May 2015 - Chiado Editora), which is considered one of the best-selling books, at national level, in the scope of the maritime-port sector. He is / was former Vice- Director of the magazine "Cluster do Mar", member of the Editorial Board of Revista de Marinha, columnist of Revista Logística Moderna and regular columnist of the main publications of the sector (Logística Hoje / Revista Mobilidade / Diário de Notícias / Newspaper "Público "/ Sector Blogs) with several dozens of articles published. He regularly participates in congresses, seminars and conferences at national and international level on the issue of Transport, Logistics, Economy, Environment, Spatial Planning and Maritime-Port Policy, organized by different entities in the sector. He is the author / co-author of dozens of communications presented within the scope of them.

Mohanad M. Hilles is from Department of Information Technology, University College of Applied Science, Palestine.

Shadi M. S. Hilles has teaching and research experience in computer science and engineering, with a Ph.D. in Computer systems and Networks. Currently, is an associate Prof, in software Engineering Department, Faculty of engineering, Istanbul OKAN University, Turkey. in addition, has 10 years academic experience in computer Science department, faculty of computer & IT, in Al-Madinah International

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

University, Malaysia. Interested area image processing, machine learning and security, Has published several research papers in indexed journals such as ISI and Scopus also book chapters in Springer and IGI, and presented in several International conferences included IEEE and was one of track chair of IEEE conference. Has participated in numerous project of computer science in courses, workshop.

Nicola Del Sarto is a post-doctoral research fellow at Scuola Superiore Sant'Anna in Pisa. He received a Ph.D in Management of Innovation, sustainability and healthcare from Scuola Superiore Sant'Anna in 2019. Nicola's research interests focus on small businesses and start-ups and support mechanisms such as incubators, accelerators and corporate accelerator programs. Moreover, he is investigating the processes of business creation under the Open Innovation paradigm. Nicola's work has been published in high-quality peer reviewed journals and presented at international conferences.

Yousef A. Baker El-Ebiary holds a Ph.D. in MIS, two Masters in IT and Business administration, a high diploma in Executive Management, and Bachelor in Software engineering. Working at Faculty of Informatics and Computing (FIK), Universiti Sultan Zainal Abidin (UNISZA), Malaysia, as a senior lecturer and member of several committees that related to postgraduate studies, conferences coordination, designing and promoting programs, this role involved in writing, design, and validation of the programs and as they developed. As well as working as a part-time senior lecturer at (MEDIU) University and (SSM) collage, Switzerland. Also Senior Fellow Researcher at (AIU) University, Malaysia. Moreover, board member of Gyancity Research Lab for research and development, India. Over 11 years experience in teaching levels of degree and graduate studies. With good experience in administrative work in the education sector as I held many positions such as dean of student affairs for 3 years, faculty's deputy for postgraduate and scientific research for 2 years, and deputy for scientific research deanship for a year. In addition, extensive experience in managing and organizing international conferences. Also supervise both Ph.D.'s and Master students, alongside working as an editor in chief for research journal and reviewer in several indexed journals as well, with experience in scientific publication and international conferences participation more than 55 international conferences with good enough number of indexed research papers (WoS and Scopus) around 100 published and in press manuscripts in fields Information Systems, Enterprise Systems, Electronic Enterprise, Internet of Things, Theory of Computation and Contemporary Management, besides some research projects locally and internationally with governments and industry. Plus, one printed book with ISBN. With a good number of awards from international exhibitions and competitions and Intellectual property has been acquired. A member in many related associations such as IEEE, IAENG, ACSE, IACSIT.

Elizaveta Fainshtein is a PhD Student at the Department of Management at the Higher School of Economics, the National Research University (HSE), Russia, Saint Petersburg. Main areas of research interests include Strategic Marketing and Business Modelling.

Lisa Ferraz holds a Master's degree in Management from the University of Aveiro.

Nelson Ferreira graduated in Electronics & Telecommunications Engineering by the University of Aveiro and pos-graduate in Business Management by the Porto Business School. In September 1995, started his functions at Bosch Termotecnologia as a software engineer in the area of microcontrollers, ERP and database design. After 2000 changed his area to the coordinator of IT area responsible for Bosch

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aveiro plant (infrastructure, server farm & support services). In 2008 initiated an internationalization experience inside Bosch, sitting in Stuttgart-Feuerbach, as responsible for administration, sizing and architecture of SAP systems for Bosch Termotechnology division worldwide. Returned to Portugal in end 2010, integrating the R&D department, in the area of electric products & systems, having during that time developed connected solutions (ZigBee, Smart Energy), namely through the participation in consortium partnership projects. In end 2014, he was named advisor of Bosch Termotecnologia SA vice-president, a function that has accumulated with the coordination of industrialization projects for new products (PMO). Starting September 2018 he has the responsibility for the coordination of Manufacturing Digitalization (Industry 4.0) for all Bosch Thermotechnology plants worldwide.

Adelaide Freitas is an Associate Professor at Department of Mathematics. She has a MSc in Probability and Statistics from University of Lisbon (1991) and a PhD in Mathematics from University of Aveiro (1998). Her scientific activity has been mainly in the area of Multivariate Statistics, Probability and Statistics in Education, Extreme Value Theory, and Statistics application in several scientific areas (e.g., social, genomic, and health). Actually, her research interests include the development and applications of sparse principal components analysis and compositional data analysis.

Ana Freitas has a Bachelor Degree in Applied Mathematics and Statistics Degree from the University of Minho and a post-graduation in Business Intelligence and Analytics from Porto Business School. Her main interests are in customer insight forecasting, predictive models and customer behaviour segmentations in the retail space at SONAE MC.

Purificación Galindo-Villardón is Professor of Statistics of the University of Salamanca. Visiting Professor in 20 European and American universities, including Trinity College, University of Oxford, and the University of Reading (England), Stanford University of California (USA), UNIBE (Costa Rica), University of Algarve (Portugal), Veracruzana University (Mexico), University of Concepción, University of Temuco, and Bernardo O´Higgins (Chile), Central University of Venezuela and University of Oriente (Venezuela), University of Bariloche (Argentina), University of Panama, ESPOL University and UNEMI University in Ecuador. Former Chair of the international PhD program in Statistics and Former Head of the Department of Statistic. Vice-Rector for Health Sciences, Quality Management and Academic Policy. University of Salamanca. Spain. Director of 64 PhD theses. More than 100 papers published in high impact scientific journals. Awarded as the "Educator of the Year 2013" in Latin America, Brazil, Spain and Portugal by the "Juarez-Lincoln-Marti" Prize from USA. Maria de Maeztu Award for Scientific Excellence in 2016.

Julia Gallardo-Pérez is a professor in the Department of Financial Economics, Accounting and Operations Management, University of Huelva, Huelva, Spain.

Petia Georgieva is Professor in Machine Learning at the Department of Electronics Telecommunications and Informatics, University of Aveiro in Portugal and senior researcher at the Institute of Electronics Engineering and Telematics of Aveiro (IEETA). She holds a Master and PhD degrees in Control System Engineering from Technical University of Sofia and a PhD degree in Computer Science from Faculty of Engineering of University of Porto in Portugal. She has held several visiting positions. In particular, she was a visiting faculty at Carnegie Mellon University (CMU)-Silicon Valley CA in 2019 (Fall Semester),

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

an invited professor at Rowan University, New Jersey, USA in 2016 (Spring Semester), a visiting faculty at CMU -Pittsburg, USA in 2012 (Fall Semester), an invited researcher at the University of Lancaster, UK in 2010 (Spring Semester). Dr. Georgieva is a Senior member of IEEE and a Senior Member of International Neural Network Society (INNS). Her research interests are in machine/deep learning and data mining with strong focus on computer vision, brain imaging, optical communications, wireless networks.

Petra Heck has been working in IT since 2002. She started as software engineer, then quality consultant and now lecturer Software Engineering. Petra has a PhD (quality of agile requirements) and started a research project on Applied Data Science and Software Engineering in February 2019. Petra holds talks on a regular basis and is author of several publications, amongst which the book (in Dutch) "Succes met de requirements".

Solène Le Bars holds PhD in cognitive neuroscience and neuropsychologist. Altran Lab, France.

Alfonso Infante Moro is a professor in the Department of Financial Economics, Accounting and Operations Management, University of Huelva, Huelva, Spain. Info: http://uajournals.com/alfonsoinfante/en/.

Juan Carlos Infante Moro is a professor in the Department of Financial Economics, Accounting and Operations Management, University of Huelva, Huelva, Spain. Info: http://uajournals.com/jcinfantemoro/en/.

Parth Patel is a ML enthusiast from the golden city of Ahmedabad. He wants to strive to make the world a better place by completing my endeavors. He has several machine learning projects that will tend towards the working of betterment of the world. One out of which is a chatbot that will make people's lives easier by having conversational command interfaces. The other is an illegal fish catching system to prevent the precious marine life of our oceans.

Riddhi Patel is a Software Engineer Intern at Gridscape Solutions, a Media Strategist at NinjaTalks, a Python Developer, a Machine Learning Enthusiast, and an Anchor.

Antoine Trad is a holder of a PhD in computer sciences degree and a DBA in business administration. He is a professor and a researcher at IBISTM in France. His research field's title is: "The Selection, Architecture, Decision Making, Controlling and Training Framework (STF) for Managers in Business Innovation and Transformation Projects"; where he published more than 60 articles on the subject. In this research project, he works on inspecting enterprise architecture solutions in business transformation projects; parallel to that he works as a consultant in enterprise architecture projects.

Alicia Izquierdo Yusta studied Economics and Business Administration in Valencia (Spain). She is assistant professor of Marketing at the University of Burgos. She received her PhD in Economics and Business Administration at the University of Burgos. She is director of the Degree in Tourism at the University of Burgos. Her research is published in international journals of business, management, and social issues such as Telematics and Informatics, Frontiers in Psychology, Service Business; European Journal of Marketing; Computers Human Behavior, Tourism Management, among others.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Norman Jacknis is currently a Senior Lecturer at Columbia University in the Technology Management Program where he teaches Machine Learning and Artificial Intelligence. He previously taught in and designed several courses for the Master's program in Applied Analytics. He is also a Senior Fellow at the Intelligent Community Forum Prior, Dr. Jacknis was Director of Cisco's IBSG Public Sector Group, where he worked with governments and businesses around the world on open innovation and technology-based strategies for marketing success. Before joining Cisco, he served more than ten years as CIO and Commissioner of Westchester County, New York government, where he was responsible for all technology and analytics, as well as technology-based economic development. Dr. Jacknis earned his BA, MA and PhD at Princeton University.

Mónica Jiménez-Hernández researcher at CIEMA (Research Center for Applied Multivariate Statistics of the University of Colima). Bachelor degree in Marketing at University of Colima and Master student in Analysis Advanced in Multivariate and Big Data at University of Salamanca, Spain. Experience in market research and digital marketing

Ana Jiménez-Zarco is Associate Professor of Innovation and Marketing Economics and Business Open University of Catalunya (UOC) and Part-Time professor at Marketing Department (ICADE-Pontificia of Comillas University). PhD in Economics and Business from the University of Castilla la Mancha and Graduate Building Models in Ecology and Natural Resource Management at the Polytechnic University of Catalonia. PhD thesis Award by the Institute of Economic Studies of Madrid. Evaluator at the European Union Programme (H2020-MSCA-ITN), is author of more than 70 articles in journals of national and international and serves as associate editor of Frontiers in Psychology Journal and Sustainability.

Özlem Karadağ Albayrak is married and has a son. She completed my primary, secondary and high school education in Istanbul. She then moved to Sakarya for university education. She graduated from Sakarya University, Department of Industrial Engineering in 1999. She worked in various companies in Istanbul for 12 years in quality, production planning, quality and production responsibility positions. She has been working as a lecturer at Kafkas University since 2013. She received her doctorate degree from Atatürk University, Department of Econometrics in 2019. She makes use of her experiences in the private sector in her lectures at the university. Later, along with my academic career, new research areas were directed to research. These areas are renewable energy sources, regional development, and logistics.

Marcel Kyas received his PhD from the University of Leiden in 2006. In his dissertation "Verifying OCL specifications of UML models", he developed compositional, computer-aided verification methods for object-oriented real-time systems. After that, he became a PostDoc at the University of Oslo. He researched type systems for dynamically evolving distributed systems. He became an assistant professor at the Freie Universität Berlin. He studied indoor positioning systems and their empirical validation. He published competitive positioning methods and an indoor positioning test-bed. Since 2015, he works at Reykjavik University. He extends his work on indoor positioning systems to distributed systems. Marcel studies statistical methods for the Internet of Things. In this context, he investigates telematics and computer security.

Markus Lamest is a PhD from Trinity College Dublin Ireland. He is interested in the area of enterprise risk management and AI governance.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Abdilahi Liban is the chief executive officer of QETC Education and Training Consultancy SDN. BHD – Malaysia, where he has been since 2018. From 2014 to 2018 he served as a part-time lecturer at Al-Madinah International University-Malaysia and Head of the Development and Training department. He received a B.IT (Hons) Information Systems Engineering. from Multimedia University - Malaysia in 2010, and a Master of Computer Science (Information Security) in 2012 from University Technology Malaysia. He received his Ph.D. in ICT from Al-Madinah International University in 2018. His research interests span both image processing and computer security. Dr. Abdilahi Liban served as an internal auditor of academic programs, university rankings and research.

Maria João Lopes is an Industry 4.0 Software Engineer at Bosch Thermotechnology, managing Industry 4.0 topics in several manufacturing plants across the world. She has a MSc in Physics Engineering with a specialization in Instrumentation from the University of Coimbra. Since 2019 she has been a student of the Doctorate in Business Innovation, a pioneer program from the University of Aveiro which aims to bring together professionals from the Industry and academia. The focus of this PhD is applying Machine Learning methods in a general way to be able to tackle different types of problems.

Sandra Maria Correia Loureiro's current research interests include relationship marketing and new technologies (VR, AR, AI), tourism marketing, and corporate social responsibility issues. She has published in top journals such as: Journal of retailing, Journal of business research, International journal of hospitality research, Journal of product and brand management, Journal of retailing and consumer services or International journal of tourism research. She recently won the 2012 Best Paper Premier Award presented by the Global Marketing Conference (comprised of EMAC, ANZMAC, KSMS, and the Japanese Association of Marketing). Highly Commended paper Award 2014 - 7th EuroMed Conference and EuroMed Research Business Institute (EMRBI), Highly Commended paper Award 2016 - 9th EuroMed Conference and EuroMed Research Business Institute (EMRBI), Best Paper Award 2016- ICCMI 2016, Highly Commended paper in the 2017 Emerald Literati Network Awards for Excellence, Best Paper Award TomiWorld 2017 for Marketing, Promotion & Consumer Behavior and Highly Commended paper in the 2018 Emerald Literati Network Awards for Excellence and 2019. Best paper award at Global Management Conference 2019 GAMMA Paris.

Mara Teresa da Silva Madaleno has a Ph.D. in Economics from the University of Aveiro (UA) since 2011, under the subject "Essays on Energy Derivatives Pricing and Financial Risk Management". Is an effective member of the Research Unit on Governance, Competitiveness and Public Policies (GOVCOPP) and Vice-coordinator of the Research line Systems for Decision Support (SAD). Currently lectures Finance and Economics at the undergraduate, graduate MSc and Ph.D. levels as an Assistant Professor at the Department of Economics, Management, Industrial Engineering and Tourism (DEGEIT) at the UA.

María Pilar Martínez-Ruiz Cathedratic professor since 2018, 21 years of research, teaching, academic and management activity at the University, and more than 16 years since obtaining the PhD degree. 6 Doctoral Theses directed and presented between 2013 and 2019 (3 with International Doctorate Mention). 103 articles published in scientific journals, 95 of which have appeared in indexed journals of relative quality. More than 90 works presented in congresses, conferences and seminars.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

João Lourenço Marques is assistant professor at the Department of Social, Political and Territorial Sciences of the University of Aveiro, lecturing courses in the areas of Urban and Regional Planning (in the field of quantitative methods and techniques to support decision making). PhD in Social Sciences by the University of Aveiro, he did a MsC in Innovation and Regional Development Policies, and his graduation studies in Urban and Regional Planning by the same university. Currently he is member of the research group on Planning and Innovation (GETIN_UA); and of the Research Unit in Governance, Competitiveness and Public Policies (GOVCOPP). He is currently coordinator of the Research Group of Systems for Decision Support – GOVCOPP; and director of the Master in Data Science for Social Sciences. His current area of interest is the development of tools and methodologies to support the decision in public policy. The application of these methodologies covers various issues (housing, demographics, mobility, planning, equipment and infrastructure, employment etc.) and various territorial levels (local to national scale).

M. Dolores Méndez Aparicio is a computer engineer and Marketing degree, PhD in Legal, Economic and Social Sciences. Since 2014, certification and active member in research at the Association for the Development of the Customer Experience, and in associations Interdisciplinary Research Group i2TIC, ACEDE and AEMARK organizations as a listener and participant in the research in the field of Customer Experience and digital behavior. Software development manager since 1982 in the Mutua Madrileña group, leader in the Spanish insurance sector. Collaborating professor at the Open University of Catalonia and reviewer of articles for prestigious magazines. As a researcher in digital client experience, she has published at various conferences, Congresses and reports. Last publication in Frontiers in Psychology.

Nigussie Mengesha is an Assistant Professor of Information Systems at the Goodman School of Business, Brock University in Canada. He teaches graduate and undergraduate courses such as Data Analysis and Business Modelling, Management of Information Systems and Technology, Database and Management of Information Systems and Technology. Prior to joining Brock University, Dr. Mengesha taught computer science and information systems courses at Ryerson University and Addis Ababa University. Dr. Nigussie Mengesha received his Ph.D. in Information Systems from the University of Oslo in Norway, and an MSc in Information Science from Addis Ababa University in Ethiopia. Dr. Mengesha's research areas include IT and Strategy, Data Analytics, Information Systems Implementation and Integration, Open Source Software, and ICT4D. He has published in Information Technologies and International Development (ITID), Electronic Journal of Information Systems in Developing Countries (EJISD), African Journal of Information Systems (AJIS), International Journal of Information Systems and Social Change (IJISSC), and major international conferences in information systems such as International Conference on Information Systems (ICIS) and International Federation of Information Processing (IFIP) WG9.4.

Filipe Miranda has a Master's Degree in Computer Degree from the University of Porto. His main areas of interest are machine learning research and data analysis. He is currently working as a data scientist on customer behaviour analysis and prediction in the retail sector.

Jorge Mora-Fernandez obtained his Ph.D. in Interactive Media, supervised by Lev Manovich and Sheldon Brown at CRCA and the VisArts Department, UCSan Diego, and by Isidro Moreno at Laboratory of Digital Culture, UCM-Spain, thanks to the European Union Social Funds and Community of Madrid

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Grants. He obtained his M.F.A. in Interactive Media and Postdoctoral specialization at IMDivision, School of Cinematic Arts, University of Southern California- USC, with La Caixa and the Fulbright Commission Grants. Jorge Mora-Fernandez is a story-writer, director, line design producer and editor ìs supervisor specialized in avant garde narratives (parallel, interactive, multi-perspectives, transmedia circular…) for documentary and fictional projects (Fictions of Reality). He is currently post producing the documentary project DQR66, Don Quixotes in Route 66. A road movie documentary about the idealist people, from different cultures and origins, living in alternative ways along the Route 66 while keeping the history of the Mother Road and its principles alive. He is a Spanish-LatinAmerican initial-mid-career cinematic artist and researcher specialized in digital storytelling who is trying to develop his creative career between USA, Europe and Latin America. He did graduate studies in California, developing later on digital narratives & multicultural cooperation projects between Latin-America-USA-Europe, such as founding the first videoart/animation LatinAmerican Festival Video Babel http://www.festivalvideobabel.org/SobreElFestivalEng.htm or the Media Literacy Workshops https://www.youtube.com/user/multiculturalvideos/videos ; mostly focus on underrepresented communities, multiculturalism, independent perspectives and recovery from addictions. He has obtained several international awards, workshops scholarship and grants as screenwriter and director of shortfilms and researcher in the areas of cinema and new media narratives: Festival Internacional de Cine de Guadalajara FICG (Between Lives Script Selected by the Program Iberoamerican Films Crossing Borders, http://www.drjorgemora.com/wp-content/uploads/2015/05/DossierEntreVidasBetweenLivesDr.JorgeMora.compressed.pdf), Festival de Cannes (Inside Track Transmedia Short-Film Awarded by the Program European Films Crossing Borders Short Film Corners https://youtu.be/j8RR-rCq1jI), Imperial Beach International Film Festival (Avant-Garde Best Short Film for Inside Track, https://youtu.be/j8RR-rCq1jI), Harold Lloyd Foundation and Electronic Arts Foundation (for his project www.multiculturalvideos.org project https://vimeo.com/129149879), an online interactive museum based on short documentaries and artistic videos, advised by Marsha Kinder Steve Anderson. Best Audiovisual Project Award by Injuve (Institute of the Youth), Spanish Ministry of Social Affairs and MediaLabMadrid presented on traveling exhibitions in all Latin-America, Short Film Corner at Cannes Festival, Festival International de Cine de Guadalajara FICG, Seoul Interactive Film Festival, International Festival of Human Rights, etc. . Inside Track was broadcasted at Canal-Arte in Europe and on La2 Spanish National-Television (https://youtu.be/j8RR-rCq1jI/https://vimeo.com/127994381). Saint Mary of Europe Best Short Film Award for his short Film Nobody, https://youtu.be/aGUI3vnBH4U, and the Imperial Beach Film Festival Award for the Best Avant Garde Short Film. He has also received grants, collaborations and recognitions from the Fulbright Commission, Siggraph, United Nations, UNESCO, AECID Spanish Agency of International Cooperation, https://youtu.be/6GCd1fLMDZI, EAP UCSD (University of California San Diego), FPI CAM EU, La Caixa, Prometheus Program, etc. His transmedia scripts were selected at Iberoamerican-Films-Crossing-Borders Program, FICG, and at European-Films-Crossing-Borders, Cannes Festival.

José Moreira received his Ph.D. in Computer Science and Networks from the École Nationale Supérieure des Télécommunications de Paris (France) and the Faculdade de Engenharia da Universidade do Porto (Portugal) in 2001. He is currently an Assistant Professor at the Department of Electronics, Telecommunications and Informatics of the Universidade de Aveiro and a researcher at IEETA, a non-profit R&D institute affiliated to the same university. His background includes programming languages, data structures and databases. His main research interests cover spatiotemporal database systems and

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Geographical Information Systems. He is also interested on spatiotemporal data analysis and knowledge engineering.

Elissa Moses is Founder & CEO, BrainGroup Global, Partner, HARK Connect and Bellwether Citizen Response, former CEO Ipsos Neuro & Behavioral Science, faculty at Columbia University on AI/ML, neuromarketing innovator, brand, advertising and technology expert. Marketing strategist with expertise in innovation, emotion and the consumer decision process . Published book, chapter and academic journal author; frequent global speaker. Advisor for boards and start-ups. Education: MS Technology Management, Columbia University and AB Human Behavior, University of Chicago.

Helena Nobre is an Assistant Professor at the Dept. of Economics, Management, Industrial Engineering and Tourism, and Researcher at the Research Unit of Governance, Competitiveness and Public Policy (GOVCOPP), at the University of Aveiro. She is an Associate Editor of the RBGN-Review of Business Management (indexed by SCOPUS Elsevier and Thomson Reuters: SSCI, JCR), and she is a member of the Editorial Board of the International Journal of Consumer Studies (Impact factor: 1.506 - ISI Journal Citation Reports © Ranking: 2018:112/147 (Business). Her main research topics are in Branding and Consumer Behavior, in particular, luxury brands, and consumer experience. Her work is published in reputable academic journals.

Rafal Ohme (PhD) is a psychologist, expert in emotions, communication and consumer neuroscience. He is a professor at WSB University in Torun and Stellenbosch University in South Africa. He held a Fulbright scholarship at Kellogg School of Management, at Stanford University he conducted research on unconscious emotions. Awarded by the President and the Prime Minister of Poland, a finalist of EY Entrepreneur of the Year 2015 program. Founder of NEUROHM, a smart tech company which licenses new research tools for HR, sales, marketing & PR. Author of articles, chapters and books including Mind Spa (2019), Emo Sapiens (2017), Unconscious Affect (2007), Subliminal Facial Information (2003), and trilogy on Automaticity (2001, 2003, 2003). Writes for Forbes.

Jan Tore Pedersen is Managing Director of Marlo a.s. He holds an MSc in Control engineering and a PhD in manufacturing from the Norwegian University of Science and Technology (NTNU) and an MSc in Computer Science from University of California at Santa Barbara. Since the early 1990ies, Dr Pedersen has been active in EU and national projects for development of innovative logistics solutions and related information technologies. Noticeable projects are e-Freight (DG MOVE) and iCargo (DG CONNECT). Since 2012, Marlo collaborated with 3M, DHL and Gebrüder Weiss to develop the Software-as-a-Service solution MixMoveMatch – aimed to ensure that transportation resources are utilised to their maximum capacity. In 2017, this development led to forming the company MIXMOVE, which is now further commercialising the MIXMOCE Solutions. Before founding Marlo, his experience includes business and technology development of advanced software systems (The Kongsberg Group), and business- and technology development in the maritime sector (The Kvaerner Group), including an integrated system for ships covering all on-board control and navigation capabilities. From 1995 to 1998, Dr Pedersen held a part-time position as professor at the Technical University of Norway in the field of mechatronics. From October 2009 until September 2011, he was an adjunct professor in logistics at the BI Norwegian Business School.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Andrea Piccaluga is Full Professor of Innovation management at the Institute of Management and delegate for Technology Transfer at Scuola Superiore Sant'Anna, Pisa. He has published several articles in leading management journals, including R&D Management Journal, Small Business Economics, California Management Review, Journal of Technology Transfer, Journal of Knowledge Management, Technology Analysis & Strategic Management, European Management Journal, and Creativity and Innovation Management. His research and teaching mainly deal with: R&D and innovation management; open innovation; and the collaboration between firms and research institutions in the innovation process.

Ilhem Quenel is in charge of the Applied AI research program at the Research and Innovation Department of Capgemini Engineering. During his PhD thesis, he worked on the application of metaheuristics (which are bio-inspired algorithms) to solve complex problems. Before joining ALTRAN 3 years ago, he was a lecturer at the University of Algiers for ten years.

Mrugendrasinh Rahevar is an Assistant professor in U &P U. Patel Department of Computer Engineering, CHARUSAT University, Gujarat, India. He received his B.E.C.E. from Gujarat University in 2006 and M.E.C.S.E. from Gujarat Technological University in 2014. Currently pursuing PhD in action recognition. His research interest includes Computer vision and deep learning.

Dorota Reykowska has over 14 years of experience in marketing research, helping build brands, generate insights, transform data into brand stories. Currently responsible for leading the R&D department in NEUROHM. Previously held the position of a Research Director in research and media agencies. Specializes in quantitative research. Expert in non-declarative studies, and application of innovative technologies to testing emotions and attitudes. Involved in the development of biometric methodology and creation of the first polish neuromarketing project. A participant and speaker on many conferences both scientific (e.g. ICORIA, NeuroPsychoEconomics), and industry (ARF, PTBRiO, Neuroconnections, etc.). A co-author of publications in polish and international journals/books.

Juan Rivera Medina has extensive work experience in the implementation of strategies and programs in public institutions linked to the Presidency of the PCM Council of Ministers; National Civil Service Authority SERVIR, National Fund for Compensation and Social Development FONCODES, Ministry of Tourism and Foreign Trade MINCETUR, Program National Agrarian Innovation PNIA, PRONAA-MIDIS National Food Assistance Program. More than 8 years of experience in advanced teaching in postgraduate degrees and three years in undergraduate degrees. Work experience with multidisciplinary, multicultural and multilingual teams in various rural regions of Peru and abroad on issues of public management, planning and foresight, strategic direction, leadership and social responsibility, conflict management and research project management in the field of social innovation.

Eugénio M. Rocha is professor of the department of mathematics and researcher at CIDMA, Univ. of Aveiro, with a PhD in Mathematics and a degree in Computer Science (artificial Intelligence), has more than 90 publications in theoretical mathematics (e.g., partial differential equations or hybrid systems) and applications into Biology, Civil Engineering, Chemistry, Econometrics, Education, and Industry. Coordinator of several projects with the Industry.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mariana Rodrigues completed her Master in Marketing at ISCTE Business School, in Lisbon. After working as a copywriter for hospitality, sports and art brands, she recently focused her professional path on content marketing for a technology company.

Deepak Saxena is an Assistant Professor at the Department of Management at Birla Institute of Technology & Science Pilani (India). Prior to this, he worked as a Research Fellow at Trinity College Dublin and as an Assistant Lecturer at Dublin Institute of Technology. He has published with the journals such as Australasian Journal of Information Systems, Journal of Information Science, Irish Journal of Management, Electronic Journal of Business Research Methods, and Electronic Journal of Information Systems Evaluations.

Gerard Schouten is a professor at the School of ICT, Fontys University of Applied Science and research fellow at the Jheronimus Academy of Data Science, both situated in The Netherlands. He holds a master's degree in physics and has a PhD in the field of cognitive science. He is a self-made deep learning and machine learning practitioner and is experienced in large-scale software development in the healthcare industry. He is member of the Methodology Advisory Board of Statistics Netherlands and the AI Advisory Broad of the province of Noord-Brabant. He initiated a new educational program on AI engineering at Fontys UAS. His research covers a wide array of topics in the broad field of AI, with a special interest in solutions that benefit society and foster our natural environment.

Nandini Sen is an anthropologist who has done her PhD (social anthropology) from Goethe University, Frankfurt, Germany, and Masters from the University of St Andrews (Social Anthropology), Scotland, UK. Her anthropological experience with the waste pickers "Urban Marginalisation in South Asia: Waste Pickers in Calcutta" was published by Routledge in 2018. Now it has been turned into a paperback edition since August 2020 after being well acclaimed by academicians and peers. Nandini had published her paper "Women and Gender in Short Stories by Rabindranath Tagore, An Anthropological Introspection on Kinship and Family' in Anthropological Journal of European Cultures in 2016. Her book chapter, How Bollywood Filmdom Operates in India and the International Arena, Consumption of Bollywood, was published in Anthropological Approaches to Understanding Consumption Patterns and Consumer Behaviour published by IGI, Global, 2020. In 2021 she has co-authored 'Gender-based Violence in India in Covid-19 lockdown'. It is published in Journal of Comparative Literature and Aesthetics, 2021. In 2020, she has published an essay, 'Pandemics, Covid-19, and literary studies, past and present' in Covid-19 Perspectives, University of Edinburgh. She has also co-authored few essays on Indian situations in Covid-19 lockdown in different academic blogs, Department of Law and Anthropology, University of Edinburgh, 2020.

Elena Serova, Assoc Prof, Ph.D. in Economics, is working now in National Research University Higher School of Economics St. Petersburg Branch. Her role combines teaching and research in equal measure. Her research interests are related to Business Analysis, Strategic Management, Marketing, and Information Management. She has co-authored books and collections of essays, regular key presenter at national and international conferences. As a research-active academic with a number of Ph.D. Students under her supervision, Elena is focusing on Complex Information Systems Modeling and Business Models in Global Environment.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Özge Sığırcı is an Assistant Professor of Marketing at Kirklareli University, Turkey. She received her Ph.D. in Marketing from Marmara University, where she also worked as a research assistant. During her Ph.D. studies, she was awarded a scholarship from the Scientific and Technological Research Council of Turkey and spent a semester at the University of Texas at Dallas as a visiting scholar. Her research interests include consumer-artificial intelligence interaction, sustainable consumption, consumer food consumption, and behavioral pricing. Her publications appeared in the Journal of Marketing and Marketing Research, Bogazici Journal, and Journal of Global Fashion Marketing. She loves playing tennis and practicing yoga.

Dora Simões received her PhD in Informatics Engineering from University of Porto, Portugal, in 2008. Actually, she is an Assistant Professor in computer science area, at the Higher Institute of Accounting and Administration – University of Aveiro (ISCA-UA), Portugal. She is member of DigiMedia Research Centre. Her current research interests include data management, organizational information systems, relationship marketing and collaborative virtual networks, and information and communication technologies in education. She has two books by IGI Global and in various conferences and journals, nationals and internationals.

Joshua Springer is a PhD student in computer science at Reykjavik University. His primary focus is control methods for autonomous, multi-rotor drones.

Nathalia Tejedor Flores is a PhD in Applied Multivariate Statistics, a full-time researcher at the Hydraulic and Hydrotechnical Research Center, professor at the Civil Engineering Faculty of the Technological University of Panama and member of the National Research System of Panamá (SNI). She is an environmental engineer with more than 5 years of experience in teaching and research, a specialist in data analysis, R-Statistical-Package, Sustainable Development, Water-Energy-Food Nexus.

Elizabete Thomas is a specialist at the National Waterway Transportation Agency, Brazil.

Adalberto Tokarski is Director at the National Waterway Transportation Agency, Brazil

Ana Maria Tomé is an Associate Professor at the Department of Electronics, Telecommunications and Informatics of the University of Aveiro, Researcher at the Laboratory of Biomedical Systems and Technologies at the Institute of Electronic Engineering and Telematics of Aveiro (IEETA). Scientific activity in digital processing, pattern recognition and machine learning for biomedical applications. The application and development of unsupervised algebraic techniques for extraction of relevant information in biomedical signals has been the focus of his activity in recent years. She has participated in several projects under the responsibility of IEETA.

Muhammad Tufail has a Ph.D. Degree in Economics from the National University of Sciences and Technology (NUST), Pakistan (Specialization in the field of Environmental and Resource Economics), a master's degree in Economics (Specialization in Technological Innovation, Economic Growth) from NUST, and a Bachelor's degree in Economics from the University of Malakand. Dr. Tufail is an Exchange Researcher at the Institute of Agronomy, University of Lisbon, and an Associate Researcher in the group of Decision Support System in the unit of Governance Competitiveness and Public Policy,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

University of Aveiro, Portugal. In his research, Dr. Tufail primarily focuses on coupling economic and ecological systems to evaluate natural resource (forest) degradation issues. While interested in many natural resources, part of his work focuses on the Himalayas and Hindu Kush conifer forest region and evaluating the economic and ecological implications of natural resources degradation to the region. He enjoys working with students and engaging students of all backgrounds and interests. He is an experienced researcher with a demonstrated history of working in the Research Industry Higher Education Commission (HEC) Pakistan, International organizations such as the International Union for Conservation of Nature (IUCN), Food and Agricultural Organization (FAO), Institute of Environmental Sciences and Engineering (IESE), National University of Sciences and Technology (NUST). Skilled in Econometric Software; (STATA, SPSS, Eviews, OxMetrics, MiniTab), Microsoft Excel, Microsoft Word, Data Analysis, Quantitative and Qualitative Research, Data Science, Report Writing, and Policy Analyst. Consultant on Assessment of Economic Valuation of Chilgoza Forest (IUCN, FAO). Consultant on Assessment of Economic Valuation of SMOGE Episodes in Pakistan (IESE, NUST, HEC).

Erika Velásquez Chacón is an economist, Master with a mention in Social Management and Sustainable Development, held positions in public and private entities for more than 20 years, an expert in development projects. Current adviser and consultant in government entities. Undergraduate teacher at the Universidad Católica San Pablo and Postgraduate at the Universidad Nacional de San Agustín de Arequipa. Speaker at conferences of the Latin American Council of Management Schools (CLADEA), in 2012 (Peru) 2013 (Brazil) 2018 (Costa Rica) 2019 (Peru) 2020 (Virtual). Speaker at the Peruvian Congress of Economic History (2019). Speaker at CAPIC (2020). Winner of the Pacific Alliance Grant - Mexico (2020). Acceptance of article and presentation at the Twentieth Ibero-American Conference on Systems, Cybernetics and Informatics CISCI (2021).

Purificación Vicente-Galindo is Professor of Statistics. Chair of the international PhD program in Statistics and Head of the Department of Statistic. University of Salamanca, Spain. Visiting Professor in European and American universities, including Trinity College, UNIBE (Costa Rica), University of Algarve (Portugal), Veracruzana University (Mexico), University of Panama, ESPOL University and UNEMI University in Ecuador. Director of 20 PhD theses. More than 40 papers published in high impact scientific journals.

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Index

A

Accuracy 7, 26, 29, 135-136, 155, 158, 162, 184, 249, 260, 272-276, 278, 282-283, 288, 299, 346-347, 350, 401, 409, 413, 431, 435, 443-444, 453, 481, 534, 552, 628, 645, 648, 650
Adjacency Matrix 310, 322
age group 66, 323, 329-330, 332
AI Aversion 342
AI Biases 342, 353, 357, 365
AI Ethical Decision Making 365
AI ethics 355, 618, 639, 641, 644, 647
AI In Business 291-292, 342, 641
AI in marketing 342-343, 345, 347, 361, 635
Algorithms 4, 24, 32, 38, 47, 103, 108, 118, 123-124, 147, 154-155, 171-172, 180-181, 208-209, 237, 243, 252, 262, 272, 277, 289, 291-292, 298, 300, 344, 350, 353-354, 358-359, 364-365, 387, 392-393, 395, 405, 443, 448-450, 453, 477, 479, 482, 487, 513, 523, 541, 545, 548, 552, 568, 571, 582-583, 585, 591, 609, 613, 620, 624, 626, 637, 641, 643, 647, 649, 651, 653
analytical AI 291, 293, 299
Analytics 3, 26, 30-31, 33, 50-51, 91, 99, 113, 115, 122, 127, 151-154, 158, 160, 162, 192, 204-205, 209, 221, 224, 230, 235, 296, 320, 358, 360, 382, 386, 391, 396, 450, 452-453, 469, 479, 481-482, 487, 522-523, 528, 531, 533-534, 536, 540-543, 546, 549, 553, 566-568, 570, 577, 585-586, 588-589, 609
Arequipeño 341
Artificial Intelligence (AI) 24, 108, 151, 159, 265, 291-292, 342, 359, 362, 365, 388, 402, 570, 598, 600-601, 614, 620, 637, 639
Artificial Intelligence Ethics 598
Asimov's Three Laws 598, 601, 609, 615, 617-618
Attention 37, 56, 119, 151-154, 160, 197, 226, 230, 234, 294, 296, 305, 329, 356, 367, 384, 408, 418, 421, 534, 536, 540, 546, 559, 586, 605, 610, 620-621, 628, 631-632, 636, 639-640
Attitudes 20, 102, 127-129, 144-146, 156-157, 163, 165, 225, 228, 343, 348, 350, 357-358, 375, 414-416, 418, 423, 426, 429, 546, 622, 638
Augmented Reality 299, 320, 570, 572, 577, 582-583, 593-595
AutoML 33-34, 263, 266, 272, 277-278, 287, 290

B

bAbi dataset 401, 409
basket 168-170, 172, 176, 178-190
Beginner Cluster 94, 102
Big Data 3, 32, 112-113, 115, 118, 121-128, 131, 135, 143-146, 157, 162, 206-207, 212, 220, 229, 239-240, 289, 344, 361, 364, 393-395, 447, 489, 507, 520, 523, 533, 540, 542-543, 547, 549, 551-554, 559-561, 563-568, 570, 607, 613, 615, 624, 637, 641-642, 645, 648-652
Bill Of Lading 8, 22, 393-395
Bottleneck 3, 270, 287, 290, 577
brand equity 367-369, 379-380
Brand Image 210, 219, 224, 226-227, 237, 239, 305, 366-369, 373-376, 378, 380, 629
Brand Perceptions 153, 367, 380
brand positioning 369, 372-374, 376, 378, 380
Brand Value 296, 367-369, 375, 380
Bulk Carrier 520
bulk shipping 381, 384-386, 391-392, 395, 397-400
Business 3-4, 6, 9, 11-12, 14, 17, 20-21, 24, 26-28, 30-32, 34, 43-44, 48-49, 52-53, 56-57, 59, 63, 79-81, 86, 88, 99, 108-115, 120-125, 127, 143, 162, 164, 166, 169, 172-177, 183, 185, 188-189, 191-194, 198-199, 201-205, 208-212, 215-216, 218-221, 223-230, 234-238, 241-243, 246-248, 260, 262, 284, 289, 291-294, 297, 300, 302-304, 306, 316-317, 319-321, 325, 328-329, 333-335, 339-341, 342, 344, 347, 358-360, 362-364, 367, 369, 374-379, 384-388, 395, 398-399, 426-427,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

429, 447-450, 452-453, 460, 462-467, 469-470, 472-475, 478-479, 481-482, 485-486, 488-490, 507, 509-511, 516-518, 521-523, 528, 533-536, 539-555, 558-559, 561, 563-570, 590, 594, 596, 609, 613, 622, 627-628, 630-634, 637, 639-642, 644-647, 649-652
Business Model 108, 113, 115, 121, 123, 125, 162, 234, 317, 334, 385, 387, 395
business modeling 223-224, 234
Business Transformation Projects 447, 481

C

Category of Products 190
Causal Map 305, 322
chartering 381-382, 384-388, 394-395, 400, 515
Chatbots 158, 202, 208, 294, 297-298, 300, 305, 345, 354, 356, 401, 404, 412, 613
CI 23, 38-40, 54
Clients 12, 43, 102, 110, 151, 158-159, 162, 194, 202, 206-207, 210, 236, 242, 250, 254, 256, 260, 300, 304, 306, 330, 545, 621
Cloud Computing 1, 14, 16-17, 109, 112, 317-320, 535, 549, 568
Cluster Analysis 85, 87, 90-92, 100-101, 159, 206, 223, 524, 531, 548
Clustering 71, 91-93, 97, 99-100, 102-103, 137, 151, 158-159, 171-172, 180-181, 186, 189-190, 239, 281, 391, 452, 477, 482, 534, 537
co-branding 366-380
Cognitive Neuroscience 165, 620, 624
Collective Cognitive Matrix 311, 322
Competitive Advantage 109, 112, 114, 116-117, 120-121, 125, 306, 368-369, 374, 517, 536, 547, 625
Consumer 3, 6, 19, 22, 143-145, 152-159, 161-163, 165, 170, 189-190, 193-194, 196-200, 205-207, 209-215, 217-218, 220-231, 234, 236-238, 261-262, 291, 294, 298, 302, 323-335, 337-341, 342-343, 345, 348-354, 356-362, 365, 369, 372, 375, 378-379, 383, 396-397, 421, 423, 427, 504, 510, 517, 545-546, 552-553, 563, 565, 619-624, 626, 628, 630-632, 634-637
Consumer Acceptance 302, 619, 631-632, 635
Consumer Algorithm Aversion 365
Consumer Behavior 3, 6, 19, 145, 152, 154, 170, 214, 222, 225, 234, 237, 294, 338, 340, 342-343, 348, 361, 396-397, 517, 553, 563, 619-620, 622, 624, 626, 637
consumer neuroscience 154, 165
Consumer Sector 291
Consumer Trends 341
Continuous Delivery (CD) 23
Continuous Integration (CI) 23
Continuous Training (CT) 39-40
Corporate Social Responsibility 55-56, 80-81
COVID-19 85-86, 126-130, 132, 134-137, 143-147, 197, 215, 250, 318, 323-325, 332-335, 337-341, 504
Critical Success Factors 124, 317, 448, 452
CRM 205, 522-524, 526, 531, 534-536, 539-540, 542
CT 39-40, 54, 487
Customer 5, 33, 59, 70, 74, 109, 111, 113, 115-120, 125, 127, 161, 163, 168-173, 175, 178-180, 182-194, 196-202, 204-228, 230-239, 265, 294, 297-301, 304-306, 313-314, 316, 334-335, 338, 345, 347-349, 353, 356-357, 359, 361-362, 367, 379-380, 412, 509, 522-524, 526, 533-543, 552, 559, 563-565, 568, 625, 628, 642, 651
customer acquisition 522-523, 526, 542
Customer Loyalty 111, 192-194, 197, 200, 213, 217, 220-221, 224-226, 235, 237, 334, 380, 533-535
customer relationship management (CRM) 522-524, 534, 542
Customer Relationship Management System 543
Customer Retention 120, 198, 221, 412, 522-523, 526, 533, 542-543
Customer Satisfaction 163, 192, 217, 219, 222, 225-228, 235, 237, 239, 294, 301, 338, 533-534, 559, 564-565, 568, 625
Customer Trust 192, 222
Customers 11, 24, 86, 108-109, 113-121, 123, 125, 151, 158-159, 168-170, 172-173, 176-181, 183-188, 193-194, 197, 199-200, 203-205, 207-208, 210-212, 220-221, 224-230, 232-240, 292, 294, 296-301, 305, 316, 328, 334-336, 343, 345-348, 354, 356-357, 367, 371, 373-374, 379-380, 392, 479, 535-536, 540, 542-543, 545, 563, 628, 631, 642, 645
Customization 168, 188, 356, 375, 586, 592

D

Data Management 20, 229, 449, 460, 472, 485, 487, 591, 619, 626-628, 630-631, 636-637
Data Mining 49, 51-52, 91-92, 100, 102, 170, 189-190, 220-221, 241-242, 261, 391, 452, 477-478, 534, 540-541, 544, 546, 549, 559, 565-566, 568, 626, 651-652
DATA MODELLING 266, 447, 452, 460, 462, 470, 472, 490
Data Science 1-2, 4-6, 14, 16-18, 24-26, 29-30, 33-35, 41-42, 44, 49-54, 126, 166, 192-194, 197-198,

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

201-202, 210, 212, 241-242, 289, 447, 449, 487-489, 506-507, 517-518, 520, 522-524, 526, 528, 534-535, 539-540, 542-550, 552-553, 555, 557-559, 561, 564-570, 573-574, 577, 582, 584-585, 589, 591-592
Decision Making 4, 8, 31-32, 123-124, 126-127, 144, 154, 163, 194, 229, 294, 316, 342, 355, 364-365, 405, 447, 485, 541, 565, 571, 635-636, 650
Decision Trees 147, 151, 158, 162, 242, 266, 272, 391, 541, 550, 552
Deep Learning 24, 31, 37, 43, 48-49, 53-54, 124, 135, 151, 157-160, 164-165, 241-242, 252, 344, 392-393, 401-402, 450, 478, 518, 550-552, 567-568, 641, 643, 647, 650-651, 653
determining factors 6, 203, 304-307, 315-316
Devops 38, 41, 49-50, 54, 536
Digital Culture 569, 573, 592
digital hub 1, 5, 16
Digital Humanities 569
Digital Migrant 323
Digital Native 323
Digital Ports 1, 4, 10-11, 16-17
Digital Ship 2, 22
Digital Supply Chain 1-2, 4, 22
Digital Technologies 112, 324-325, 331, 340-341, 520, 542, 563, 566, 611
Digital Transformation 6, 20, 110, 112-113, 115, 122-123, 125, 193, 210-211, 220, 289, 318, 381, 385, 388, 395, 507
dynamic panel data analysis 491, 494, 497-500, 502

E

E-Commerce 190, 210, 217, 224, 226-227, 240, 318, 322-323, 326, 328, 333-334, 336-337, 339, 548, 552-554, 640
Econometrics 144, 501-502, 544-546, 548-550, 559, 565-566
Economic Growth 57, 120, 328, 491-494, 497-501, 504, 612
Economics 20-21, 80, 122-124, 137, 146, 213, 223, 236, 239, 303, 317, 320, 326, 334, 337, 340, 364, 396-398, 478, 503, 544-553, 555, 557-559, 561, 564-568, 634, 648
EDTV 431, 435, 438-439, 443-444
Education 4.0 506, 508, 519
EEA Grants 381, 386, 396, 400, 506, 516, 520
e-government development 85-89, 92-93, 95-98, 100, 102-103
E-Government Development Index (EGDI) 85, 89, 102
E-government Evolution 85
Electrodermal Activity (EDA) 154, 165
Electroencephalography (EEG) 154, 165
Electronic Government (E-Government) 102
Emergent cluster 93-94, 102
Emerging Technologies 110, 120, 297, 382, 388-389, 398, 551
Emotion 144, 151, 153-155, 163, 165, 203, 216, 219, 345, 489, 606, 621, 632-634, 637-638
Empathetic AI 293, 299, 609
energy production 492, 501, 503-504
Enhancement 123, 431, 434-437, 441-445, 513, 638, 646, 650
ENIDH 506, 515, 518, 520
ENIDH (Escola Superior Náutica Infante D. Henrique) 520
Ensemble Learning 147, 542, 568
enterprise architecture 447, 474, 490
EWOM (Electronic Word-of-Mouth) 239
explainable AI 639, 645-646, 648-651, 653
External Logistics Biplot 55-56
Eye Tracking 155, 619, 621-622, 625-626, 628, 635, 637

F

Facial Action Unit Coding (FAC) 165
Facial Coding 155, 621-622, 625, 638
fingerprint 431-432, 434-440, 442-446
Fmri 151, 155-156, 165, 622, 624, 626
fNIRS 151, 154, 156-157, 163, 165
freight rate 397, 400
Functional Magnetic Resonance Imaging (fMRI) 155, 165
Functional Near-Infrared Spectroscopy (fNIRS) 154, 156, 165
fuzzy cognitive maps 304, 307, 309-310, 317-318, 321

G

Game Design 582-583, 590, 596-597
Game Design Document 583, 597
Game Engine 597
gameplay 579-580, 582, 586, 589-590, 597
Generational Groups 330, 341
Gestural Coding 625, 638
Global Fortune 500 Companies 55, 74, 81
Global Reporting Initiative 55-56, 58, 78-80
Globalization 19-20, 114, 116, 192, 222, 327, 367, 396, 503, 612
Graph 72, 263, 266-269, 271-273, 282-283, 285-288, 290, 322, 409, 553, 557, 571, 576
Growth Mindset 507, 514, 518, 520

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

H

Heatmap 585, 597
HJ-Biplot 55, 57, 60-62, 64, 73, 77-78, 81
Hospitality 220, 227, 236-239, 291-292, 294, 297, 300, 302, 318, 345-346, 356, 366-370, 374, 376-377, 379, 559, 566, 568, 651
hotel brand 366-367, 369-370, 372-373, 375
Hotels 294, 296-297, 304-307, 309-310, 315-316, 346, 369, 372, 374-378, 539
human-AI hybrid 639-640, 643-646, 653
Hybrid AI 345, 365

I

iCode 128-129, 147
iCode™ 129, 147
ICT 32, 87, 91-93, 99-100, 164, 193, 304, 319-321, 327, 329, 331, 333, 338, 340, 385
Image 16, 77, 110, 125, 147, 152, 161, 192, 199, 210, 214, 219, 224, 226-227, 237, 239, 264, 305, 344-345, 366-369, 371-378, 380, 392, 404, 431, 434-435, 437-441, 443-445, 450, 552, 565, 574, 579, 616, 624, 629, 631
IMO Number 22
Implementation 10, 21, 27, 33, 35, 88, 92, 118, 122, 172, 180, 202, 242, 250, 281, 298, 304, 306, 316, 318-320, 324, 333-334, 343, 371, 384-385, 387, 401-402, 404, 407, 411-412, 447-448, 456, 469-470, 478-479, 484, 486-487, 493-494, 500, 509-512, 514, 517-518, 534-535, 546, 564-565, 570, 576, 588, 590, 601, 626, 628, 630, 637, 644, 651
Industry 4.0 112, 121, 123-124, 263, 381, 385, 398, 506-507, 559, 566
informal sector 324, 341
Information Technology (IT) 32, 86, 103
Innovation 3, 5, 10-12, 20-22, 52, 110-111, 114, 117-118, 121-125, 164, 192-193, 202, 210-211, 214, 219, 221-222, 227, 238, 282, 289, 294, 302, 305-306, 317, 319-321, 333, 350, 357, 360, 379, 382, 398-399, 494, 506-507, 509-511, 516-521, 559, 617, 630
interactive media 569, 572-574, 576, 586, 591-592
Interactive Storytelling 569, 572, 574, 577, 588, 593, 595
intuitive AI 293, 299

K

Knowledge Management 100, 122, 190, 218, 229, 238, 508, 516, 519
knowledge transfer 506, 508-509, 516-517, 520

L

Latent 61-62, 78, 206, 364, 431-432, 434-440, 442-446, 533, 586
Learning Process 161, 330, 353, 507-508, 640-643, 653
Likert Type Scale 414-415, 417, 429
listing product 173, 182-184, 190
Living Lab 506, 511, 517-520
Living Labs 506, 511-512, 515-516, 519-521
Logistics Biplot 55-56, 81
Logistics Single Window (JUL – Janela Única Logística) 22
Long Short-Term Memory (LSTM) 170, 241-243, 262, 412
luxury 366-379, 624

M

Machine Learning 2, 4, 6, 17, 20, 23-24, 28, 42, 48-54, 85, 87, 91-92, 98, 102-103, 109, 111, 113, 115, 117-119, 122-123, 125-128, 132, 135, 144, 146-147, 151-152, 157-165, 189, 209, 213, 221, 224-225, 230, 237, 241-242, 260-263, 265, 270, 272, 277, 282, 285, 288-290, 293-294, 296, 298, 300, 344, 353, 361-362, 382, 391, 393, 396, 402, 405, 407, 411-412, 450, 533-535, 544-545, 549-553, 565, 567-568, 583, 589, 599, 635, 639-645, 647-653
Machine Learning (ML) Engineering 23
machine learning techniques 85, 87, 91, 103, 117, 119, 126-128, 164
Manufacturing Execution System (MES) 290
Manufacturing Production Line 263, 267, 278
maritime economy 1, 9, 13
Maritime Education 512-515, 517-518, 520
Maritime Industry 5, 381-383, 395, 399, 512
Maritime Logistics 1-2, 4, 6, 20, 397
maritime supply chain 1, 4-5, 21
Marketing 21, 59, 100, 110, 123, 125, 153, 156-158, 160, 162-164, 166, 168-169, 188-189, 192-193, 197, 199, 201, 205-206, 210-222, 224-230, 234-240, 301-303, 305-306, 320, 323, 326, 329, 333-334, 336, 341, 342-345, 347-348, 356-364, 367-369, 372-373, 377-379, 414, 418, 423-427, 497, 503, 522-523, 534-535, 539, 542-543, 546, 548, 552-553, 559, 561, 564-566, 581, 592, 608, 619-620, 622, 626, 629, 631-638, 651
Marketing 5.0 192, 211, 222
marketing channels 323, 326, 329, 333-334, 336, 341

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Marketing Research 100, 189, 213, 229-230, 302, 361, 364, 377, 414, 418, 423-424, 426-427, 620, 622, 629, 632
mathematical models 447, 449, 487-488, 490, 545-546
means of payment 323, 326, 330, 332
mechanical AI 291, 293, 299
Memory 151-152, 170, 180, 196-199, 207, 212, 215, 241-245, 262, 288, 343, 401-402, 406-408, 411-412, 464, 475, 552, 567, 620, 622, 636, 638
Memory Networks 401-402, 406-407, 411-412, 567
meta-analyses 151-152
ML 23-26, 28-49, 52-54, 118, 151-158, 160, 162-163, 242, 252, 265-266, 272, 274, 277, 281-285, 287, 402, 639, 641-648
ML engineering 25, 28, 33, 41, 44-48, 53-54
ML system 24-26, 28-29, 31-32, 35-37, 40-41, 46, 54
MLOps 23, 25, 38-42, 46-47, 49-50, 53-54
Models 3, 6, 17, 24, 26, 28-29, 31-32, 34-35, 37-40, 44, 46, 50, 52-54, 59, 86, 88, 109, 111, 120-123, 132, 135-136, 144, 154-157, 159-162, 170, 193, 196, 198, 202, 206-207, 209-210, 221, 225, 242-243, 246, 252, 254, 260-262, 265, 267, 272, 284, 294, 317, 321, 334, 347, 350, 384, 393, 399, 405, 407, 409, 447, 449-450, 453, 460, 463-464, 468-469, 471, 475, 479, 482, 486-490, 493, 498-499, 501, 509-510, 531, 533, 535-536, 545-547, 552, 561, 565, 567-568, 572, 576, 595, 611, 624-625, 633, 640-641, 643, 645-646, 648, 652-653
moral development 639, 644, 647-648
multivariate analysis 62, 64, 79-81, 224, 227, 229-231, 234-235, 239
Multivariate Analysis of Variance (MANOVA) 239
Multivariate Sentiment Analysis 223

N

native basket 168-169, 172, 176, 178, 181-188, 190
Natural Language Processing (NLP) 156, 209, 393, 413
Network 3, 5-6, 9-11, 15, 18, 117, 145, 156, 158-159, 161-162, 165, 212, 227, 237, 241-243, 252-254, 259-262, 307, 332, 334, 392, 397, 402, 406, 408, 411-413, 435, 445, 492, 513, 523, 536-537, 539-540, 549, 552-554, 558, 568, 570, 643, 647, 653
Neural Networks 24, 31, 37, 113, 115, 135, 157-159, 170, 207, 215, 241-243, 260-262, 266, 344, 359, 392-393, 402, 404, 406, 413, 476, 478, 545, 550, 552, 641, 648
Neuroethics 619, 629, 633, 638, 650
neuromarketing 2.0 619-622, 624-632, 634, 638
Neuroscience 151-154, 156-157, 163, 165, 326, 616, 619-620, 622, 624, 626, 633-634, 638
Non-Parametic Test 429
Non-parametric Test 414

O

Open Innovation 20, 506-507, 509-510, 517-521
Optimization 1, 5, 14-16, 156-157, 252, 261-262, 266, 272, 277, 282, 287, 289, 387, 392, 397-399, 533, 544, 546, 552, 568, 570, 579, 592, 642

P

panel data 491, 494-495, 497-504, 546
Paperless Port 8, 22
Paperless Port (PsP – Porto sem Papel) 22
Parallel Coordinates 55, 57, 62, 64, 73, 76, 78, 81
parameter stability 263, 270-272
Parametric Test 414, 420, 423, 429
People 3, 5, 9, 43-44, 57, 75, 88, 100, 108-109, 111, 113, 115-118, 120, 127-129, 131-132, 134-135, 137, 143, 156-159, 161, 196, 201, 203, 267, 282, 284, 293, 307, 321, 323-327, 329-334, 340-341, 346, 348, 350-355, 357, 359, 362, 388, 390, 395, 403-404, 481, 506-508, 520-521, 543, 546, 571-572, 579, 592, 599-600, 602-603, 608-609, 611-614, 616, 618, 626-631, 643, 645-646, 651
Personalization of Experience 304
Pilot Project 506
Pipeline 23, 32, 35, 40, 42, 54, 272-278, 281-282, 284-285, 288-290, 452, 534
Population 56-57, 90, 130, 137, 155-157, 161, 193, 247, 278, 281, 307, 323-327, 331-335, 337, 341, 383, 404, 491-495, 497-500, 504
Port Authority 8, 22, 388
port network 15
Prediction 61, 70, 123, 130, 136, 144-145, 147, 152, 154-156, 158, 163-164, 168, 170, 172, 180, 189, 237, 242, 255-256, 259, 261, 263-266, 270-273, 277-278, 281-283, 285, 288-289, 402, 408, 478, 535-536, 541-542, 550-552, 567, 641, 643, 648, 650, 653
Predictive Analytics 33, 154, 205, 452, 469, 479, 531, 542, 546, 567
Predictive Model of Consumer Behavior 222
predictive robustness 263, 271-272
price index 168-173, 183-191
Privacy by Design 639, 644-645, 648-649, 651-653
privacy concerns 342, 352-353, 361, 626, 634, 636
Processes 2, 4-6, 9, 11-12, 15-16, 18, 23, 33, 40, 43, 48, 86-88, 108-113, 115, 117-118, 120, 122, 124, 127, 145, 152-155, 157, 165, 200, 206-209, 211-

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

212, 221, 229, 241, 244, 264, 266, 282, 284-285, 288-289, 292, 294, 297-299, 301-302, 305-306, 324, 326, 328, 333, 335-337, 341, 346, 348, 381-382, 386, 388, 393, 395, 408, 450, 453, 457, 460, 463-466, 471, 476-479, 481-482, 484-485, 490, 508, 510, 514, 517, 520, 522-523, 526, 535, 542-543, 546, 553-554, 568, 577, 585, 590-591, 596, 600, 611, 619-620, 629, 649
Processing Time Distribution 263
production environment 23, 26-27, 35-36, 39, 54
Products 28, 42, 54, 58-60, 70, 73, 108-109, 113-115, 117-120, 125, 147, 151, 155-156, 158, 161, 163, 168-184, 186-188, 190-191, 193, 201, 203, 206-207, 210-211, 224-230, 239-240, 242, 261, 263-264, 267, 271, 287, 292, 294, 298-299, 302, 306, 323-326, 329, 333-336, 343, 346, 350-353, 356-357, 367, 369, 372-376, 380, 383, 400, 509-510, 542, 563, 569-570, 586, 590-591, 601, 620, 624, 626, 628
Progressor cluster 93, 95, 103
promotional activity 169-170, 180, 183
Prophet 241-242, 246, 248, 250, 254-256, 258-260, 262
Public Investment 493, 504
Public Port 7, 22
Pupil Tracking 625, 638

Q

qualitative data 366, 368, 370, 375
question answering system 405, 408, 412-413
Question Answering Systems 401, 405
Queue Directed Graph 263, 267, 269

R

Random Forest 126-128, 135-136, 143-144, 146-147, 272, 277-278, 535
Recurrent Neural Networks 170, 260, 262, 393, 406, 413
Regional Development Degrees 504
regional economic growth 491, 493, 498, 500
regional population 491, 495, 497-498, 500
regional public investments, 491, 500
regional renewable energy 491, 495, 497-498, 500
Reneeable Energy 491
Renewable Energy 12, 491-495, 497-504, 552
renewable energy production 501, 503-504
Representative Basket 168
response time (RT) 128-129, 147
Retail 72, 78, 168-170, 177, 180, 183, 188-190, 221, 241-243, 250, 260, 262, 291-292, 298-301, 333-334, 337-340, 367, 377, 536
Retailing 213, 219-221, 291-292, 345, 362
Robots 110, 236, 294, 296, 298, 300, 303, 305, 344, 346-347, 356, 361, 363-364, 559, 565, 568, 598-599, 601-610, 612-618, 631

S

Saas 44, 381, 387, 400, 536
SaaS (Software as a Service) 400
Segmentation Analysis 101
Sentiment Analysis 110, 155-156, 158, 163, 165, 223, 230-231, 234, 237, 239, 523, 552
Sentiment Analysis of the Text (Direction of the Review) 239
service format 223, 226, 228, 230, 234, 239
Shipping 4.0 381, 385, 396, 398
Shipping Industry 2, 381-385, 388-390, 393, 395-396, 400, 512, 515
shopping mission 171, 178-180, 182, 184-185, 191
Single Window 10, 20-22
Single Window Environment 10
Small Data 206-207, 212, 216, 222, 391
smart port 382
Smes 108-110, 112-117, 119-122, 124-125, 283, 320, 377
Software Engineering 23-24, 26, 28-36, 39, 42-44, 48-54, 318, 320, 478, 488-489, 540, 650
Spillover Effects 368, 380
stability period 184, 187, 191
Statistics for Decision Support 544
stock-keeping unit 168, 170, 172-173, 191
Strategic Partnership 366
Strong AI 299, 344-345, 365, 641
substitution effect 145, 177, 188, 190-191
Supervised 24, 29, 34, 147, 158, 165, 252, 391, 406-407, 477, 641, 653
Supervised Learning (SL) 165
Supply Chain 1-5, 8, 11-12, 14, 16, 18, 20-22, 67, 71, 111, 117-118, 122, 266, 289, 322, 337, 528, 542, 564, 566
sustainability gradients 55-56, 62, 70, 75, 78
Sustainability Report 81
Systematic Literature Review 43, 122, 519-520, 522-525, 531, 539-540, 542-543, 548

T

TEU – Twenty-Foot Equivalent Unit 22
Text analytics 26, 158
Text Mining 204, 223-224, 228-230, 234, 238-240, 362, 452, 523, 531, 533, 537, 540-542, 565-566, 568

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use

Text Mining Technologies 240
time series forecasting 241-242, 260-262
token 267-271, 283, 287-288, 290
Tourism 88, 212-214, 218, 221, 236-239, 291-292, 297-298, 300-301, 303, 317-319, 346, 364, 377, 487, 539, 542, 551-553, 559, 561, 564-568, 587
TRADf 448, 454, 456-459, 462, 469, 474, 483-484, 487, 490
Turing's Test 598, 618

U

Unsupervised Learning 158, 165
unsupervised machine learning 85, 87, 91, 98, 103
User Experience 201, 210, 519, 569, 572, 574-575, 577, 579-580, 582, 584-585, 589, 591, 596

V

vaccination 126-132, 134-137, 141, 143-145
Value Proposition 183, 223-228, 230, 232-237, 240, 369, 524, 537
Vanishing Gradient Problem 243, 413
Veteran cluster 93-95, 103
video games 569-572, 579-580, 582-583, 586, 590-591, 593-596
Virtual Assistants 112, 119, 345-347, 360
virtual product 173, 191
Virtual Reality 155, 164, 193, 215, 569-570, 573, 580, 586, 593-594, 596
Vocal Coding 625, 638
Voice of the Customer 197-198, 222

W

Weak AI 299, 344-345, 365
wine brands 366-367, 369-376

EBSCOhost - printed on 2/9/2023 8:10 AM via . All use subject to https://www.ebsco.com/terms-of-use