

Premier Reference Source

# Responsible AI and Ethical Issues for Businesses and Governments



Bistra Vassileva and Moti Zwilling

**IGI Global**  
PUBLISHER OF TIMELY KNOWLEDGE

# Responsible AI and Ethical Issues for Businesses and Governments

Bistra Vassileva  
*University of Economics, Varna, Bulgaria*

Moti Zwilling  
*Ariel University, Israel*

A volume in the Advances in Human  
and Social Aspects of Technology  
(AHSAT) Book Series



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](mailto: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: Vassileva, Bistra, 1966- editor. | Zwilling, Moti, 1969- editor.

Title: Responsible AI and ethical issues for businesses and governments /

Bistra Vassileva and Moti Zwilling, editors.

Other titles: Responsible artificial intelligence and ethical issues for businesses and governments

Description: Hershey, PA : Engineering Science Reference, [2021] | Includes bibliographical references and index. | Summary: "This book is aimed at scholars and practitioners who want to widen their understanding of artificial intelligence out of the 'narrow' technical perspective to a more broad viewpoint that embraces the links between AI theory, practice, and policy"-- Provided by publisher.

Identifiers: LCCN 2020021023 (print) | LCCN 2020021024 (ebook) | ISBN 9781799842859 (hardcover) | ISBN 9781799864387 (paperback) | ISBN 9781799842866 (ebook)

Subjects: LCSH: Artificial intelligence--Moral and ethical aspects. |

Artificial intelligence--Industrial applications.

Classification: LCC Q334.7 .R47 2021 (print) | LCC Q334.7 (ebook) | DDC 174/.4--dc23

LC record available at <https://lccn.loc.gov/2020021023>

LC ebook record available at <https://lccn.loc.gov/2020021024>

This book is published in the IGI Global book series Advances in Human and Social Aspects of Technology (AHSAT) (ISSN: 2328-1316; eISSN: 2328-1324)

#### 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](mailto:eresources@igi-global.com).



# Advances in Human and Social Aspects of Technology (AHSAT) Book Series

ISSN:2328-1316  
EISSN:2328-1324

Editor-in-Chief: Ashish Dwivedi, The University of Hull, UK

## MISSION

In recent years, the societal impact of technology has been noted as we become increasingly more connected and are presented with more digital tools and devices. With the popularity of digital devices such as cell phones and tablets, it is crucial to consider the implications of our digital dependence and the presence of technology in our everyday lives.

The **Advances in Human and Social Aspects of Technology (AHSAT) Book Series** seeks to explore the ways in which society and human beings have been affected by technology and how the technological revolution has changed the way we conduct our lives as well as our behavior. The AHSAT book series aims to publish the most cutting-edge research on human behavior and interaction with technology and the ways in which the digital age is changing society.

## COVERAGE

- Gender and Technology
- Cyber Bullying
- Computer-Mediated Communication
- Digital Identity
- Human Rights and Digitization
- Information ethics
- Technology Adoption
- End-User Computing
- Cultural Influence of ICTs
- Philosophy of technology

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](mailto:Acquisitions@igi-global.com) or visit: <http://www.igi-global.com/publish/>.

The Advances in Human and Social Aspects of Technology (AHSAT) Book Series (ISSN 2328-1316) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, [www.igi-global.com](http://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-human-social-aspects-technology/37145>. 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.



## Titles in this Series

*For a list of additional titles in this series, please visit:*

<http://www.igi-global.com/book-series/advances-human-social-aspects-technology/37145>

### ***Machine Law, Ethics, and Morality in the Age of Artificial Intelligence***

Steven John Thompson (University of California, Davis, USA & University of Maryland Global Campus, USA)

Engineering Science Reference • © 2021 • 266pp • H/C (ISBN: 9781799848943) • US \$295.00

### ***Bridging the Gap Between AI, Cognitive Science, and Narratology With Narrative Generation***

Takashi Ogata (Iwate Prefectural University, Japan) and Jumpei Ono (Aomori University, Japan)

Information Science Reference • © 2021 • 409pp • H/C (ISBN: 9781799848646) • US \$190.00

### ***Reductive Model of the Conscious Mind***

Wieslaw Galus (Independent Researcher, Poland) and Janusz Starzyk (Ohio University, USA)

Information Science Reference • © 2021 • 296pp • H/C (ISBN: 9781799856535) • US \$195.00

### ***Dyslexia and Accessibility in the Modern Era Emerging Research and Opportunities***

Kamila Balharová (Pedagogical and Psychological Counseling Center, Brno, Czech Republic)  
Jakub Balhar (Gisat s.r.o., Czech Republic) and Věra Vojtová (Masaryk University, Czech Republic)

Information Science Reference • © 2021 • 279pp • H/C (ISBN: 9781799842675) • US \$165.00

### ***Understanding the Role of Artificial Intelligence and Its Future Social Impact***

Salim Sheikh (Saïd Business School, University of Oxford, UK)

Engineering Science Reference • © 2021 • 284pp • H/C (ISBN: 9781799846079) • US \$200.00

*For an entire list of titles in this series, please visit:*

<http://www.igi-global.com/book-series/advances-human-social-aspects-technology/37145>



701 East Chocolate Avenue, Hershey, PA 17033, USA

Tel: 717-533-8845 x100 • Fax: 717-533-8661

E-Mail: [cust@igi-global.com](mailto:cust@igi-global.com) • [www.igi-global.com](http://www.igi-global.com)

# Table of Contents

**Preface**..... xiv

**Acknowledgment**..... xix

**Section 1**  
**Philosophical and Conceptual Foundations of AI**

**Chapter 1**  
Artificial Intelligence: Concepts and Notions ..... 1  
*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

**Chapter 2**  
The Soul of Artificial Intelligence and Races’ Separation of AI and Homo ..... 19  
*Rinat Galiautdinov, Independent Researcher, Italy*

**Chapter 3**  
Practical Issues in Human and Artificial Intelligence Interaction ..... 35  
*Arthur Kordon, Kordon Consulting LLC, USA*

**Chapter 4**  
Policy and Management Issues of Artificial Intelligence ..... 54  
*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

**Section 2**  
**Responsible Application of AI Tools and Methods**

**Chapter 5**  
What We Should Have Learned From Cybersyn: An Epistemological View  
on the Socialist Approach of Cybersyn in Respective of Industry 4.0 ..... 68  
*Dietmar Koering, Arphenotype, Germany*

## **Chapter 6**

The Transformation and Enterprise Architecture Framework: The Applied Holistic Mathematical Model for Geopolitical Analysis (AHMM4GA).....80

*Antoine Trad, Independent Researcher, Croatia*

## **Chapter 7**

A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge, and Behavior Regarding Cyber Security ..... 128

*Galit Klein, Ariel University, Israel*

*Moti Zwilling, Ariel University, Israel*

*Dušan Lesjak, International School for Social and Business Studies, Slovenia*

## **Chapter 8**

Developing Cyber Security Competences Through Simulation-Based Learning ..... 148

*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

## **Section 3**

### **Responsible and Ethical Considerations for AI Deployment**

## **Chapter 9**

The Influence of COVID-19 Outbreak on the Readiness of Firms to Cyber Threats..... 165

*Moti Zwilling, Ariel University, Israel*

## **Chapter 10**

Ethical Aspects of Information Literacy in Artificial Intelligence ..... 179

*Selma Leticia Capinzaiki Ottonicar, Sao Paulo State University (UNESP), Brazil*

*Ilídio Lobato Ernesto Manhique, Sao Paulo State University (UNESP), Brazil*

*Elaine Mosconi, Université de Sherbrooke (UdeS), Canada*

## **Chapter 11**

Artificial Intelligence and Ethical Marketing .....202

*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

*Plamena Palamarova, University of Economics, Varna, Bulgaria*

## **Chapter 12**

Liability in Labor Legislation: New Challenges Related to the Use of Artificial Intelligence .....	214
--	-----

*Andriyana Andreeva, University of Economics, Varna, Bulgaria*

*Galina Yolova, University of Economics, Varna, Bulgaria*

<b>Compilation of References .....</b>	<b>233</b>
--	------------

<b>About the Contributors .....</b>	<b>255</b>
-------------------------------------	------------

<b>Index.....</b>	<b>258</b>
-------------------	------------

# Detailed Table of Contents

**Preface**..... xiv

**Acknowledgment**..... xix

**Section 1**  
**Philosophical and Conceptual Foundations of AI**

**Chapter 1**

Artificial Intelligence: Concepts and Notions ..... 1

*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

In recent years, artificial intelligence (AI) has gained attention from policymakers, universities, researchers, companies and businesses, media, and the wide public. The growing importance and relevance of artificial intelligence (AI) to humanity is undisputed: AI assistants and recommendations, for instance, are increasingly embedded in our daily lives. The chapter starts with a critical review on AI definitions since terms such as “artificial intelligence,” “machine learning,” and “data science” are often used interchangeably, yet they are not the same. The first section begins with AI capabilities and AI research clusters. Basic categorisation of AI is presented as well. The increasing societal relevance of AI and its rising inburst in our daily lives though sometimes controversial are discussed in second section. The chapter ends with conclusions and recommendations aimed at future development of AI in a responsible manner.

**Chapter 2**

The Soul of Artificial Intelligence and Races’ Separation of AI and Homo .....19

*Rinat Galiautdinov, Independent Researcher, Italy*

Artificial intelligence breaks into our lives. However, some questions are already being raised, and increasingly, these issues affect aspects of morality and ethics. Is it possible to scoff at thinking AI? When will it be invented? What prevents us from

writing laws of robotics right now, putting morality into them? What surprises does machine learning bring us now? Can machine learning be fooled, and how difficult is it? But even greater questions arise in the context of the separation of races of AI and humans. Is AI racism our future? This chapter explores these questions.

**Chapter 3**

Practical Issues in Human and Artificial Intelligence Interaction .....35

*Arthur Kordon, Kordon Consulting LLC, USA*

The chapter will focus on some practical issues in human and AI interaction based on the experience of applying AI in several large corporations. The following issues will be discussed: weaknesses of human intelligence, weaknesses of AI, benefits of human intelligence from AI, negative effects of AI on human intelligence, resistance of human intelligence toward AI, and how to improve the interaction between human and artificial intelligence. The discussed issues will be illustrated with examples from real-world applications.

**Chapter 4**

Policy and Management Issues of Artificial Intelligence .....54

*Bistra Konstantinova Vassileva, University of Economics, Varna,  
Bulgaria*

The capacity for AI research, technology, and application is seen as vital to national competitiveness, security, and economic strength. In the last few years, several countries and regions have developed and released AI strategic plans, thus setting up a race to become the global leader in the field. The chapter starts with an overview of the latest development in AI legislation and governance principles. The first section begins with a review of available policies and strategies on AI by countries and regions. Some best practices in AI governance are presented as well. The specifics of AI ecosystems are discussed in the second section. Gephi software tool is used to visualize the mapping of the Italian AI ecosystem. The chapter ends with conclusions and recommendations aimed at the future development of policy and management for responsible AI implementation.

**Section 2**

**Responsible Application of AI Tools and Methods**

**Chapter 5**

What We Should Have Learned From Cybersyn: An Epistemological View  
on the Socialist Approach of Cybersyn in Respective of Industry 4.0 .....68

*Dietmar Koering, Arphenotype, Germany*

Currently, a major topic is what changes will digitalization and the fourth industrial revolution bring to our society. It is clear that digital transformation of society and

the introduction of new technologies will make many jobs obsolete. This process logically leads to the idea of a universal basic income (UBI). In this respect, the socialist project, Cybersyn, is of great interest because it constituted a prototype of a data- and people-related idea to solve this problem. The aim was to increase the country's production, while counteracting rising unemployment through a socialist paradigm, which is obviously pertinent to the development of Industry 4.0. Although Cybersyn can be considered as an early prototype and catalyst, today's exponentially greater computational power has made such systems real, and humans are often excluded from them. Human beings are also positively affected by digital transformation. Herein, the current work contributes to the ethical debate concerning the digital transformation of society.

**Chapter 6**

The Transformation and Enterprise Architecture Framework: The Applied Holistic Mathematical Model for Geopolitical Analysis (AHMM4GA).....80  
*Antoine Trad, Independent Researcher, Croatia*

This chapter proposes the applied holistic mathematical model for geopolitical analysis (AHMM4GA) that is the result of research on societal, business/financial, and geopolitical transformations using applied mathematical models. This research is based on an authentic and proprietary mixed research method that is supported by an underlining mainly qualitative holistic reasoning model module that punctually calls to quantitate functions. The proposed AHMM4GA formalism, attempts to simulate functions to support empirical processes.

**Chapter 7**

A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge, and Behavior Regarding Cyber Security .....128  
*Galit Klein, Ariel University, Israel*  
*Moti Zwilling, Ariel University, Israel*  
*Dušan Lesjak, International School for Social and Business Studies, Slovenia*

With the COVID-19 pandemic, many organizations and institutions moved to e-learning and to e-working from home. With the increase in internet usage, the rate of cyber-attacks have also increased, and this was followed by the request for more cyber security behaviors from employees and students. In the current study, the authors explore the connection between cyber security awareness, cyber knowledge, and cyber security behavior. The authors measured the behaviors among students in two similar countries: Israel and Slovenia. Results show that students felt they had adequate awareness on cyber threat but apply only a few protective measures to protect their devices, usually relatively common and simple ones. The study findings also show that awareness to cyber threats mediate the connection between knowledge



and protection behaviors, but only in the case that the knowledge is specific with regard to IT protection courses. Results, implications, and recommendations for effective cyber security training programs for organizations and academic institutions are presented and discussed.

**Chapter 8**

Developing Cyber Security Competences Through Simulation-Based Learning ..... 148

*Bistra Konstantinova Vassileva, University of Economics, Varna, Bulgaria*

The importance of cyber security competences is growing both in practice and in academia during the last few years. This chapter provides a current overview of the existing body of the literature in the field of simulation-based learning and the key cyber security issues. The author’s primary goal is to develop a methodological business-oriented and evidence-based learning framework which will provide students or trainees with the opportunity to develop practical skills in the field of cyber security issues through a virtual business simulator. The overall intention is to provide a coherent framework that makes use of active-based learning and gamification to support the active participation of students or trainees. To meet these goals, the Reference Framework for Applied Competences (REFRAC) is applied. Taking into account that in 2040 ICT and internet will be ‘culturally invisible’, cyber security competences will be a must for everyone. They will be critical both for personal and companies’ survival in the turbulent and highly competitive digital environment.

**Section 3**

**Responsible and Ethical Considerations for AI Deployment**

**Chapter 9**

The Influence of COVID-19 Outbreak on the Readiness of Firms to Cyber Threats..... 165

*Moti Zwilling, Ariel University, Israel*

Technology impacted the lives of millions of people with their home day-to-day activities. When the COVID-19 pandemic struck in many countries, there was a need to change both the mode of working with technology as well as to handle internet and online risks exposure. During the pandemic, cybercrime groups utilized the internet usage to commit cybercrimes especially by exploiting vulnerabilities of many applications, networks, and infrastructures. This study aims to explore the impact of COVID-19 on the readiness of organizations to handle cyber threats in two directions: 1) analysis of CVE common vulnerability data before and during the pandemic period and 2) analysis of fuzzy logic data model designed to demonstrate the importance of firms readiness to cope with cyber threats. Results show that due

to the significant increase in cyber threats, small firms tend to be more fragile to cyber threats than big ones, and they have to invest more resources to mitigate cyber threats. Findings and implications are discussed.

**Chapter 10**

Ethical Aspects of Information Literacy in Artificial Intelligence .....179

*Selma Leticia Capinzaiki Ottonicar, Sao Paulo State University  
(UNESP), Brazil*

*Ilídio Lobato Ernesto Manhique, Sao Paulo State University (UNESP),  
Brazil*

*Elaine Mosconi, Université de Sherbrooke (UdeS), Canada*

The purpose is to analyse information literacy to provide ethical insight into artificial intelligence. The methodology was based on a systematic literature review of SCOPUS, Web of Science, Library and Information Science Abstracts, and Science Direct. The results demonstrated that there are only a few studies about the topic, so there is a research opportunity about this type of literacy and its ethical aspects in the context of artificial intelligence. As a conclusion, information literacy is crucial to the development of critical thinking in technology use. Information literacy should be applied in artificial intelligence courses to discuss ethical aspects of technology.

**Chapter 11**

Artificial Intelligence and Ethical Marketing .....202

*Bistra Konstantinova Vassileva, University of Economics, Varna,  
Bulgaria*

*Plamena Palamarova, University of Economics, Varna, Bulgaria*

In this chapter, the author argues that technologies will transform the marketing organization and reshape the marketing activities of companies. The aim of the chapter is to summarize the main challenges of digital disruption as well as to identify their implications to the legal and ethical aspects of digital and interactive marketing activities. The research aims driving this chapter are related to the identification of the main challenges regarding the legal protection of social media customers. Survey results about social media behaviour and cybersecurity issues are presented and discussed.

## **Chapter 12**

### **Liability in Labor Legislation: New Challenges Related to the Use of Artificial Intelligence .....214**

*Andriyana Andreeva, University of Economics, Varna, Bulgaria*

*Galina Yolova, University of Economics, Varna, Bulgaria*

The study analyzes the influence of artificial intelligence on labor relations and the related need to adapt to the legal institute of liability in labor law with the new social realities. The sources at European level are studied and the current aspects of liability in the labor law at a national level are analyzed. Based on the analysis, the challenges are outlined and the trends for the doctrine, the European community, and the legislation for the introduction of a regulatory framework are identified.

### **Compilation of References ..... 233**

### **About the Contributors ..... 255**

### **Index..... 258**

# Preface

Theoretical and applied field of Artificial Intelligence (AI) is vast and quite diverse. From theoretical perspective AI encompasses logic, probability, continuous mathematics, perception, reasoning, and learning. Different scientific approaches are related to AI such as cybernetics, statistical learning, brain simulation, etc. The AI tools and techniques (machine learning, data mining, advanced analytics, neural networks) are evolving in a high speed. It is incredibly difficult for the practitioners and even for the researchers to keep an eye on the latest developments in the AI area especially from the applied perspective, especially the application of AI technology in science and industry. Following the rise of AI from technological perspective the ethical issues of AI application in different industry sectors and life aspects are gaining increased attention. Expanding development of artificial intelligence poses a myriad of questions regarding the responsible AI development (in terms of AI methods and applications) and its implementation. The prevailing notion is that AI should be accountable, explainable, transparent, and fair for all organizations and individuals.

The objective of this book is to provide the basic philosophical and conceptual foundations of AI as well as to draw the attention of the readers on the responsible application of AI tools and methods. It will be written for professionals who want to improve their knowledge and skills on responsible and ethical AI implementation in different industries.

The primary target audience of this book are scholars and practitioners who want to widen their understanding of AI out of the ‘narrow’ technical perspective to a more broader viewpoint which embraces the links between AI theory, practice, and policy. The secondary target audience are students who are interested in applied side of AI.

Chapter 1, “Artificial Intelligence: Concepts and Notions,” starts with a critical review on AI definitions since terms such as “artificial intelligence,” “machine learning,” and “data science” are often used interchangeably, yet they are not the same. The first section begins with AI capabilities and AI research clusters. Basic categorisation of AI is presented as well. The increasing societal relevance of AI and its rising inburst in our daily lives though sometimes controversial are discussed

## **Preface**

in second section. The chapter ends with conclusions and recommendations aimed at future development of AI in a responsible manner.

Chapter 2, “The Soul of Artificial Intelligence and Races’ Separation of AI and Homo,” focuses on the effects of AI on our lives. However, some questions are already being raised, and increasingly, these issues affect aspects of morality and ethics. Is it possible to scoff at thinking AI? When will it be invented? What prevents us from writing laws of robotics right now, putting morality into them? What surprises does machine learning bring us now? Can machine learning be fooled, and how difficult is it? But even greater questions arise in the context of the separation of races of AI and humans. Is AI racism our future? The author of the chapter provide an interesting discussion on these questions.

Chapter 3, “Practical Issues in Human and Artificial Intelligence Interaction,” focuses on some practical issues in human and AI interaction based on the experience of applying AI in several large corporations. The following issues are discussed: (i) weaknesses of human intelligence, (ii) weaknesses of AI, (iii) benefits of human intelligence from AI, (iv) negative effects of AI on human intelligence, (v) resistance of human intelligence toward AI, (vi) how to improve the interaction between human and artificial intelligence. The discussed issues are illustrated with examples from real-world applications.

Since the capacity for AI research, technology, and application are seen as vital to national competitiveness, security, and economic strength, Chapter 4, “Policy and Management Issues of Artificial Intelligence,” provides an overview on developed and released AI strategic plans on national level. The chapter starts with an overview of the latest development in AI legislation and governance principles. The first section begins with a review of available policies and strategies on AI by countries and regions. Some best practices in AI governance are presented as well. The specifics of AI ecosystems are discussed in the second section. Gephi software tool is used to visualize the mapping of the Italian AI ecosystem. The chapter ends with conclusions and recommendations aimed at the future development of policy and management for responsible AI implementation.

The major topic of Chapter 5, “What We Should Have Learned From Cybersyn: An Epistemological View on the Socialist Approach of Cybersyn in Respective of Industry 4.0,” is about changes which digitalization and the fourth industrial revolution will bring to our society. It is clear that digital transformation of society and the introduction of new technologies will make many jobs obsolete. This process logically leads to the idea of a universal basic income (UBI). In this respect, the socialist project, Cybersyn, is of great interest, because it constituted a prototype of a data- and people-related idea to solve this problem. The aim was to increase the country’s production, while counteracting rising unemployment through a socialist paradigm, which is obviously pertinent to the development of Industry 4.0.

Although Cybersyn can be considered as an early prototype and catalyst, today's exponentially greater computational power has made such systems real, and humans are often excluded from them. Human beings are also positively affected by digital transformation. Herein, the current work contributes to the ethical debate concerning the digital transformation of society.

Chapter 6, "The Transformation and Enterprise Architecture Framework," proposes the Applied Holistic Mathematical Model for Geopolitical Analysis (AHMM4GA) that is the result of research on societal, business/financial and geopolitical transformations using applied mathematical models. This research is based on an authentic and proprietary mixed research method that is supported by an underlining mainly qualitative holistic reasoning model module that punctually calls to quantitate functions. The proposed AHMM4GA formalism, attempts to simulate functions to support empirical processes.

With the COVID-19 pandemic many organizations and institutions moved to e-learning and to e-working from-home. With the increase in internet usage so is the rate of cyber-attacks, that followed by the request for more cyber security behaviors from employees and students increased. Chapter 7, "A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge, and Behavior Regarding Cyber Security," provides a study which aim is to explore the connection between cyber security awareness, cyber knowledge and cyber security behavior. The authors measured the behaviors among students in two similar countries: Israel and Slovenia. Results show that students felt they had adequate awareness on cyber threat but apply only few protective measures to protect their devices. The study findings also show that awareness to cyber threats mediate the connection between knowledge and protection behaviors, but only in the case that the knowledge is specific and regard to IT protection courses. Results, implications, and recommendations for effective based cyber security training programs for organizations and academic institutions are presented and discussed.

Since the importance of cyber security competences is growing both in practice and in academia during the last few years, Chapter 8, "Developing Cyber Security Competences Through Simulation-Based Learning," provides a current overview of the existing body of the literature in the field of simulation-based learning and the key cyber security issues. The author's primary goal is to develop a methodological business-oriented and evidence-based learning framework which will provide students or trainees with the opportunity to develop practical skills in the field of cyber security issues through a virtual business simulator. The overall intention is to provide a coherent framework that makes use of active-based learning and gamification to support the active participation of students or trainees. To meet these goals, the Reference Framework for Applied Competences (REFRAC) is applied. The author concludes that cyber security competences are a must for everyone nowadays. They

## ***Preface***

are critical both for personal and companies' survival in the turbulent and highly competitive digital environment.

When the COVID-19 pandemic stroke in many countries, there was a need to change both the mode of working with technology as well as to handle internet and online risks exposure. During the pandemic, cybercrime groups utilized the internet usage to commit cybercrimes especially by exploiting vulnerabilities of many applications, networks and infrastructures. Chapter 9, "The Influence of COVID-19 Outbreak on the Readiness of Firms to Cyber Threats," aims to explore the impact of COVID-19 on the readiness of organizations to handle cyber threats in two directions: 1. Analysis of CVE common vulnerability data before and during the pandemic period; 2. Analysis of fuzzy logic data model designed to demonstrate the importance of firms readiness to cope with cyber threats. Results show that due to the significant increase in cyber threats, small firms are tend to be more fragile to cyber threats than big ones and they have to invest more resources to mitigate cyber threats.

The purpose of Chapter 10, "Ethical Aspects of Information Literacy in Artificial Intelligence," is to analyse information literacy to provide ethical insight into artificial intelligence. The methodology was based on a systematic literature review of SCOPUS, Web of Science, Library and Information Science Abstracts and Science Direct. The results demonstrated that there are only a few researches about the topic, so there is a research opportunity about this type of literacy and its ethical aspects in the context of artificial intelligence. As a conclusion, information literacy is crucial to the development of critical thinking in technology use. Information literacy should be applied in artificial intelligence courses to discuss ethical aspects of technology.

In Chapter 11, "Artificial Intelligence and Ethical Marketing," the authors argue that technologies will transform the marketing organization and reshape the marketing activities of companies. The aim of the chapter is to summarize the main challenges of digital disruption as well as to identify their implications to the legal and ethical aspects of digital and interactive marketing activities. The research aims driving this chapter are related to the identification of the main challenges regarding the legal protection of social media customers. Survey results about social media behaviour and cybersecurity issues are presented and discussed.

The study presented in Chapter 12, "Liability in Labor Legislation: New Challenges Related to the Use of Artificial Intelligence," analyzes the influence of artificial intelligence on labor relations and the related need to adapt to the legal institute of liability in labor law with the new social realities. The sources at European level are studied and the current aspects of liability in the labor law at a national level are analyzed. Based on the analysis, the challenges are outlined and the trends for the doctrine, the European community and the legislation for the introduction of a regulatory framework are identified.



The evolution of AI bursts with contradictions. AI possesses a huge potential to improve human lives, and, at the same time, it could widen the social and digital divides. In order to utilize its positive potential and to minimize the threats, it is critical to involve people (experts, scientists, policy makers, funders and investors, etc.) with various background and organizations from different industries to build a solid background of common language and shared understanding of AI capabilities and risks to guide all stakeholders to positive impacts. A broader engagement of civil society on the values that need to be embedded in AI and the directions for future development are also needed. The key to success is to balance the transformational potential of AI with human safety and privacy.

Despite the increasing adoption of AI tools and applications there are still some limitations which are pure technical, practical limitations and limitations in use. Future research directions will focus these general limitations. Regarding responsible AI, future research areas should cover the following topics: bias, transparency and fairness in AI algorithms; responsible AI operationalization, geopolitical impacts of AI.

The AI benefits are obvious but there are certain societal risks related to the diffusion of AI technologies in products and services which requires an open debate about AI governance. The main focus should be placed on developing an internationally recognised ethical and legal framework for the design, production and application of AI. This framework should be based on common AI principles and should provide a roadmap for protecting humanity by responsible uses of AI technologies. Unless AI is still at a relatively early stage of development and large scale industrial applications are yet to be developed, the societal challenges of AI applications should be explored and prioritized especially within the context of AI ecosystem.

# Acknowledgment

The editors of this book are very grateful to the contributing authors who did their best to accomplish this challenging book project. The scientific discussions about responsible and ethical AI were fascinating and inspiring. We respect the comments and advices of peer reviewers which motivated us to think critically.

We appreciate the encouragement and support of our colleagues and friends.

Without these valuable contributions, completion of this book would not have been possible. Thank you very much!


Section 1

# Philosophical and Conceptual Foundations of AI

# Chapter 1

## Artificial Intelligence: Concepts and Notions

**Bistra Konstantinova Vassileva**

 <https://orcid.org/0000-0002-5976-6807>  
*University of Economics, Varna, Bulgaria*

### ABSTRACT

*In recent years, artificial intelligence (AI) has gained attention from policymakers, universities, researchers, companies and businesses, media, and the wide public. The growing importance and relevance of artificial intelligence (AI) to humanity is undisputed: AI assistants and recommendations, for instance, are increasingly embedded in our daily lives. The chapter starts with a critical review on AI definitions since terms such as “artificial intelligence,” “machine learning,” and “data science” are often used interchangeably, yet they are not the same. The first section begins with AI capabilities and AI research clusters. Basic categorisation of AI is presented as well. The increasing societal relevance of AI and its rising inburst in our daily lives though sometimes controversial are discussed in second section. The chapter ends with conclusions and recommendations aimed at future development of AI in a responsible manner.*

### INTRODUCTION

Artificial Intelligence is a powerful and transformative technology. On one hand, AI embraces huge potential to provide real social, economic and environmental benefits. It is considered as an area of critical importance for national competitiveness. McKinsey Global Research, for instance, estimates that the use of AI will add as much as \$13 trillion to global GDP by 2030 (Bughin et al., 2018). On the other

DOI: 10.4018/978-1-7998-4285-9.ch001

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

hand, the growing use of algorithms raises potential ethical concerns toward the societal relevance of AI.

There are many inconsistencies and confronting points of view when it comes to AI. Artificial intelligence quite often is compared and even is defined related or opposed to human intelligence. Regarding its application AI is divided into “strong” (AI for general applications) and “weak” AI (AI for practical applications) thus contradicting cognitive science versus engineering.

The chapter starts with a critical review on AI definitions since terms such as “artificial intelligence,” “machine learning,” and “data science” are often used interchangeably, yet they are not the same. The first section begins with AI capabilities and AI impact. Basic categorisation of AI is presented as well. The increasing societal relevance of AI and its rising penetration in our daily lives though sometimes controversial are discussed in second section. The chapter ends with conclusions and recommendations aimed at future development of AI in a responsible manner.

## **Background**

In 1956, American computer scientist John McCarthy organised the Dartmouth Conference, at which the term ‘Artificial Intelligence’ was first adopted. Research centres emerged across the United States to explore the potential of AI. If we assume the following stages of the life cycle of technology: 1/ technological invention or discovery, 2/ technological emergence, 3/ technological acceptance, 4/ technological sublime (the value of technology is fully appreciated), and 5/ technological surplus (Kendall, 1999) where exactly could we position AI nowadays? Is AI an emerging technology?

The emergence of almost every new technology is accompanied by estimates of the speed with which it will obtain universal use. In fact, an examination of such cases indicates that the adoption cycle of the new technology - from invention through early use to more widespread use and then general use - is usually about 25 years (Ein-Dor, 2011). According to Anyoha (2020) the adoption cycle in AI will be much longer. This research shows, for instance, that Arthur Clarke and Steve Kubrik prediction that “...by the year 2001 we will have machines with intelligence that matched or exceeded human’s” was quite overstated.

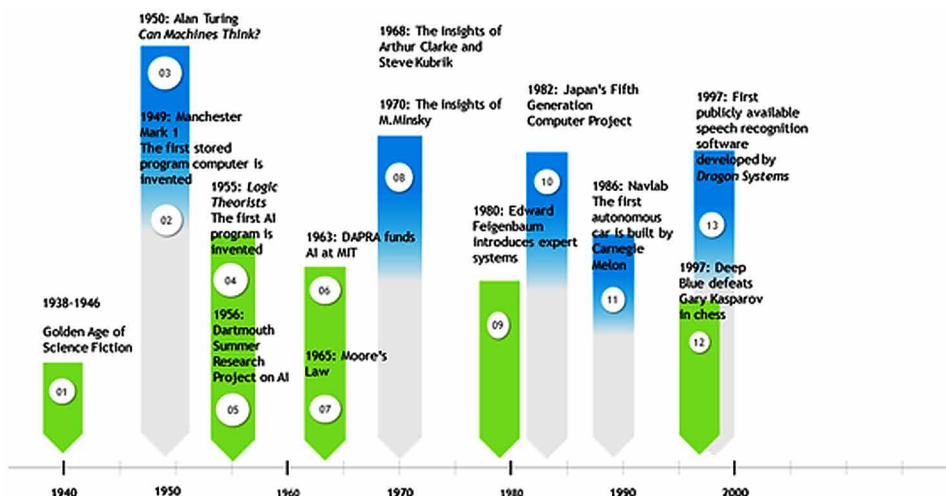
The evolution timeline of AI is presented in Figure 1.

As it shown in the figure, AI is still being invented 50 years after the idea first evolved. Various AI technologies is still being created, even the conceptual foundations of AI are continually evolving. Thus, AI is still very much an emerging technology. The Gartner (2019) predictions regarding the life cycle of AI main tools and applications are shown in Table 1.

## Artificial Intelligence

Figure 1. The evolution timeline of AI

Source: Adapted by Anyoha, R., 2020. *The History Of Artificial Intelligence - Science In The News*. [online] *Science in the News*. Available at: <<http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>> [Accessed 24 September 2020].



AI lacks a universal definition (Olley, 2020:8). In the broadest terms, AI refers to the creation of machines (agents) that think and act like humans. Common definitions of AI focus on automation and, as a result, often fail to make clear the opportunities available to IT and business leaders (Panetta, 2019). In order to reveal the hidden opportunities of AI for greater personalization and differentiation, a focused approach on AI application to augment human decision making and interactions is needed. It is important to identify how technology is applied during critical interactions with customers and then to apply AI to those points for additional business value.

Certain confusion of terms is evident which demands clarification especially when it is necessary to ensure that policy objectives are correctly translated into research priorities, that education matches job market and business needs, and comparison between various countries and regions across the globe could be done in a reliable way. The apparent lack of a common language across perspectives calls into question the quality of understanding and communication across the AI field. A common language and understanding would better connect actors in the AI ecosystem.

Russell and Norvig (2010:2) outlined four approaches to AI definition along the following two dimensions: type of thinking (humanly and rationally) and type of acting (humanly and rationally) (Figure 2). Their analysis shows that all four approaches to AI have been followed during the years, each by different people with different methods.

Table 1. AI life cycle predictions

Time frame to reach the plateau of productivity	AI tool / application
Less than 2 years	<ul style="list-style-type: none"> <li>■ Robotic process automation software</li> <li>■ Speech recognition</li> <li>■ GPU accelerators</li> </ul>
2 to 5 years	<ul style="list-style-type: none"> <li>■ Augmented intelligence</li> <li>■ Data labeling and annotation services</li> <li>■ AI related C&amp;SI services</li> <li>■ AI developer toolkits</li> <li>■ Edge AI</li> <li>■ Deep neural network ASICs</li> <li>■ Intelligent applications</li> <li>■ AutoML</li> <li>■ Chatbots</li> <li>■ Deep neural networks (Deep learning)</li> <li>■ Machine learning</li> <li>■ VPA-enabled wireless speakers</li> <li>■ FPGA accelerators</li> <li>■ Virtual assistants</li> <li>■ Computer vision</li> <li>■ Insight engines</li> </ul>
5 to 10 years	<ul style="list-style-type: none"> <li>■ AI marketplaces</li> <li>■ Reinforcement learning</li> <li>■ AI governance</li> <li>■ Neuromorphic hardware</li> <li>■ Decision intelligence</li> <li>■ AI cloud services</li> <li>■ Knowledge graphics</li> <li>■ Explainable AI</li> <li>■ Smart robotics</li> <li>■ AI PaaS</li> <li>■ Digital ethics</li> <li>■ Conversational user interfaces</li> <li>■ Graph analytics</li> <li>■ NLP</li> <li>■ Cognitive computing</li> </ul>
More than 10 years	<ul style="list-style-type: none"> <li>■ Artificial general intelligence</li> <li>■ Quantum computing</li> <li>■ Autonomous vehicles</li> </ul>

Source: Adapted by Gartner, <https://www.enterpriseirregulars.com/144131/whats-new-in-gartners-hype-cycle-for-ai-2019/>

Substantial progress in many areas has accelerated the development of AI, which has the potential to reshape the competitive landscape of companies, jobs, and the economic development of countries (Bughin et al., 2018:5). But the key challenge for such development is the speed of adoption of AI which could widen gaps between countries, companies, and workers reinforcing the current digital divide (van Dijk, 2012). The potentials of AI are rooted mainly in widespread prospective applicability



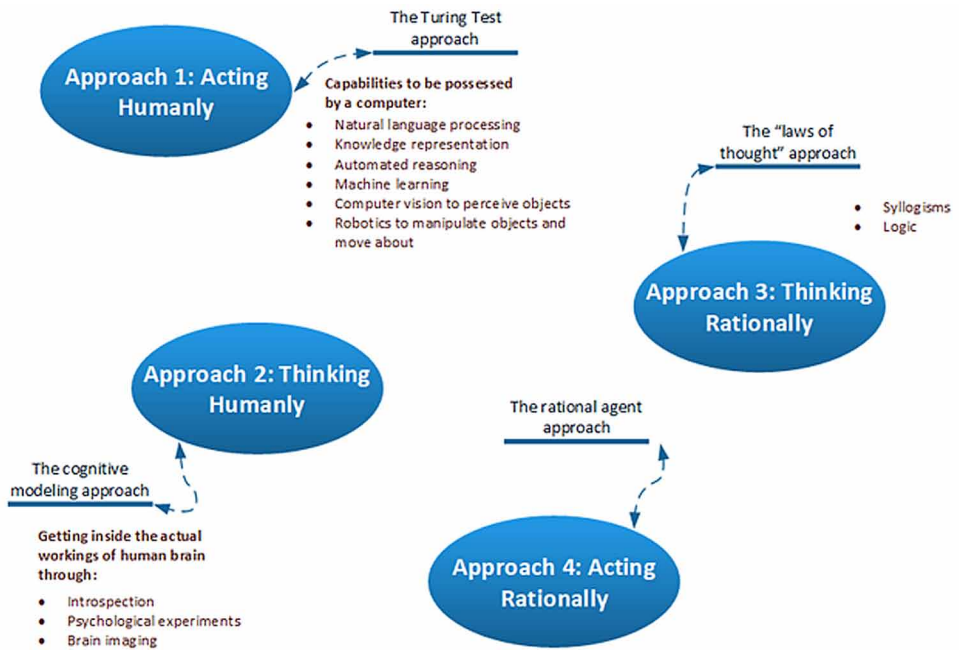
Artificial Intelligence

Table 2. Summary of AI definitions

Author	Definition
Chui and Malhotra (2018)	The ability of a machine to perform cognitive functions associated with human minds (such as perceiving, reasoning, learning, and problem solving), includes a range of capabilities that enable AI to solve business problems.
Panetta (2019)	AI is applying advanced analysis and logic-based techniques, including machine learning, to interpret events, support and automate decisions, and take action. AI is technology that emulates human performance, typically by learning from it.
Nilsson (2010)	Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment.
Russell and Norvig (2010:30)	The intelligence is concerned mainly with rational action. Ideally, an intelligent agent takes the best possible action in a situation.
Copeland (2019)	Artificial intelligence is the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

Source: Author’s development

Figure 2. The main AI approaches  
Source: Adapted by Russell and Norvig (2010:4-5)



through the entire business system and improvement on human biases (Chui and Manyika, 2018) while the limitations are caused by the current AI life stage which is characterized with a very rapidly evolving set of techniques and technologies which are going through development.

## **MAIN FOCUS OF THE CHAPTER**

### **AI Capabilities and Categorization**

The field of artificial intelligence is broad, dynamic, and promptly evolving. The key output of this process are technologies with enormous global societal implications (WEF, 2020; Schwab, 2016; Hager et al., 2017). The value to be created annually across the entire economy by AI and AI-based technologies is estimated to reach trillions of dollars (Chui and Malhotra, 2018).

According to the Gartner analysis (2020) the following three trends will affect AI in the next few years. First, natural language processing, generation and contextual interpretation will improve the communication both ways with people. Second, deeper and broader integration of AI with existing applications and IoT projects which will drive business and service value. Third, as AI adoption rate increases and it becomes more common, it will work effectively with others employing similar technologies thus making ecosystem interaction richer.

As it was mention above, AI impact on economy is growing but it differs significantly between industries. Five segments of industries could be identified based on the AI impact (billion USD) and share of AI impact in total impact (%).

Retail sector is a single outperformer with AI impact of more than 600 billion USD and share of AI impact above 40%.

The results from Global AI Survey, conducted by McKinsey (Cam et al., 2019:4) indicate growth in AI adoption in nearly every industry in 2019. Retail has seen the largest increase while high tech leads in AI adoption. Moreover, industries are generally using the AI capabilities most relevant to their value chains.

The most common AI business applications nowadays are presented in Table 4.

Transportation, healthcare, education, low-resource communities, public safety and security, employment and workplace, home/service robots, and entertainment are considered the domains where AI is already having or is projected to have the greatest impact.

The AI applications are grounded in the AI capabilities which include robotic process automation, machine learning, conversational interfaces, computer vision, natural language (NL) text understanding, NL speech understanding, NL generation, physical robotics, autonomous vehicles. According to Elsevier AI Report 2020

*Table 3. Segmentation of industries based on AI impact*

Segment	Industries	AI impact, billion USD	Share of AI impact in total impact
Low impact – Low share	<ul style="list-style-type: none"> <li>■ Telecommunications</li> <li>■ Agriculture</li> <li>■ Pharmaceuticals and medical products</li> <li>■ Aerospace and defense</li> <li>■ Chemicals</li> <li>■ Media and entertainment</li> </ul>	Below 200	Up to 40
Medium impact – High share	<ul style="list-style-type: none"> <li>■ High tech</li> <li>■ Oil and gas</li> </ul>	Arround 200	From 40 to 50
Medium impact – Low share	<ul style="list-style-type: none"> <li>■ Insurance</li> <li>■ Banking</li> <li>■ Basic materials</li> <li>■ Advanced electronics/semiconductors</li> </ul>	From 200 to 300	Up to 40
High impact – Low share	<ul style="list-style-type: none"> <li>■ Healthcare systems and services</li> <li>■ Consumer packaged goods</li> <li>■ Public and social sectors</li> </ul>	Above 300	Up to 40
High impact – High share	<ul style="list-style-type: none"> <li>■ Transport and logistics</li> <li>■ Travel</li> <li>■ Automotive and assembly</li> </ul>	Above 300	From 40 to 60

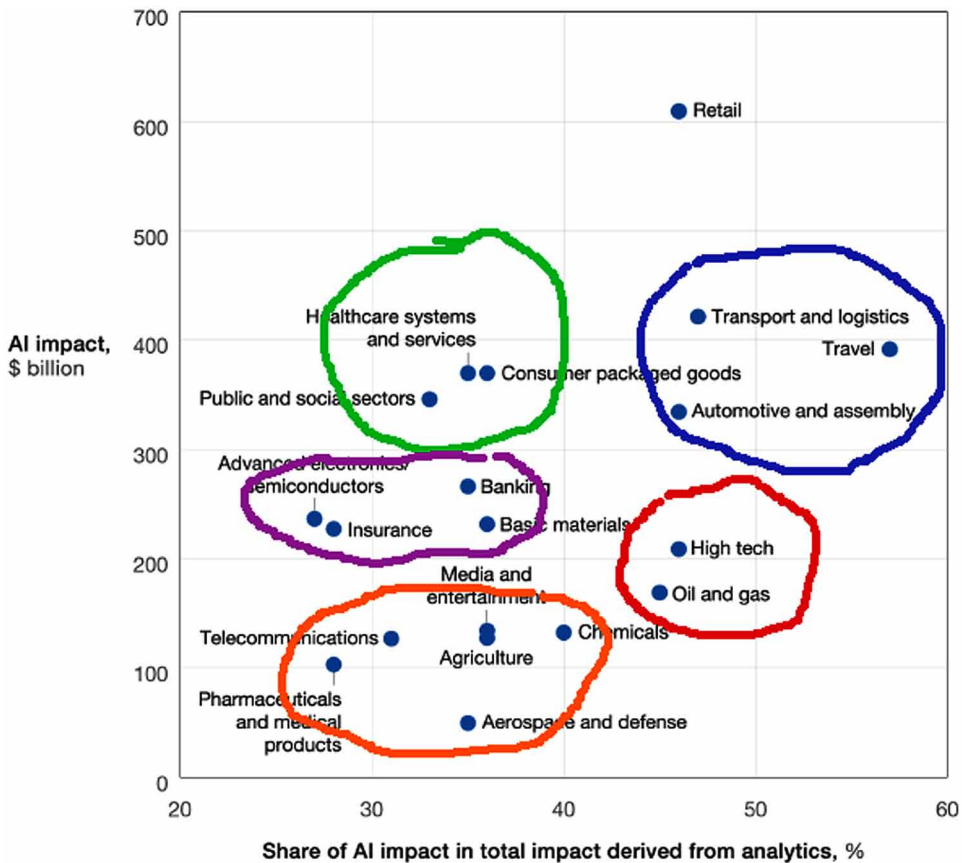
Source: Adapted using the data provided by Chui, M. and Malhotra, S. (2018). AI Adoption Advances, But Foundational Barriers Remain. McKinsey Analytics. [online] McKinsey & Company. Available at: <<https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain>> [Accessed 28 April 2020].

(Olley, 2020) the field and scope of recent AI research include the following main topics: search and optimization, fuzzy systems, natural language processing (NLP) and knowledge representation, computer vision, machine learning and probabilistic reasoning, planning and decision making, and neural networks. The popularity of the above mentioned AI applications and techniques was changing over the AI life cycle. In the late 90s neural nets and genetic algorithms were the two most popular techniques with a number of specific application areas such as human-machine interfaces, expert systems, autonomous intelligent machines (AIM), and data mining (Ein-Dor, 1999). Nowadays, advances in facial and speech recognition have produced virtual assistant technologies that are being integrated into our daily life. AI enables companies to collect data from a wide variety of places and apply self-improving analysis.

AI could be categorized in several aspects depending on different criteria. First, based on the level of intelligence of the AI application, they could be referred as low-intelligence applications (automation) and higher-end intelligence (decision-making). Second, AI applications are divided on centrally controlled and distributed

*Figure 3. Industry segments by AI impact*

Source: Adapted by Chui, M. and Malhotra, S. (2018). *AI Adoption Advances, But Foundational Barriers Remain*. McKinsey Analytics. [online] McKinsey & Company. Available at: <<https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain>> [Accessed 28 April 2020].



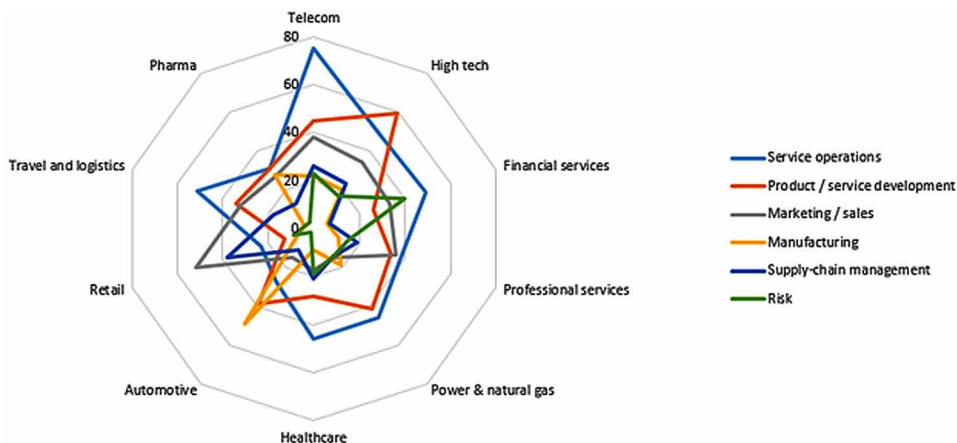
among multiple machines regarding the type of control. Third, depending on the level of sophistication the following types of AI are defined:

- Reactors (low level of individual entity intelligence);
- Categorizers (low-to-medium level of individual entity intelligence);
- Responders (medium level of individual entity intelligence);
- Learners (medium-to-high level of individual entity intelligence);
- Creators (high level of individual entity intelligence).

## Artificial Intelligence

*Figure 4. AI adoption by industries and by business functions*

Source: Data were gathered from the following McKinsey reports: <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence>; <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence>.



*Table 4. The most common AI business applications*

Industry sector	Brief description / example
Sales and marketing	Customization of the sales process Personalization of communications to prospects and clients Matching sales staff to buyers Personalized pricing AI-based recommender systems have revolutionized online search optimization and digital ad targeting.
Service	Virtual customer assistance and triage Prediction of maintenance and upcoming repair needs Connecting service staff to customers Discovering process gaps
Supply chain	Discovery and correction of data errors Discovering risks in the supply chain Elevating insights from Internet of Things (IoT) devices in the field Logistics planning
Banking and financial services	Helping customers access their bank balances using chatbots.
Healthcare	Follow up with patients post-discharge using virtual nursing assistants. Improving medical image analysis for rapid and accurate diagnoses and treatment planning.

Source: Adapted by Andrews, W., 2020. Build The AI Business Case. [online] Emtemp.gcom.cloud.  
Available at: <<https://emtemp.gcom.cloud/ngw/globalassets/en/information-technology/documents/trends/ai-business-case-ebook.pdf>> [Accessed 28 April 2020].

Fourth, a differentiation could be done between weak AI, i.e., machines that can simulate thinking within a narrow context to accomplish a specific task, and strong AI, i.e., intelligent machines that can reason. Strong AI is often based on observations of the processes of human intelligence, and from this point of view, any successful attempt to build a strong AI becomes a theory of the way humans perform such activities (Ein-Dor, 1999). Weak AI has much more restricted goals and may be viewed as attempts to induce computers to behave at the level of intelligent humans on specific, predefined tasks. Fifth, depending on the scope of the tasks to be completed the AI techniques are divided into narrow vs. general AI. Narrow AI describes an AI that is limited to a single task or a set number of tasks. General AI describes an AI which can be used to complete a wide range of tasks in a wide range of environments. NITI Aayog (2018:15) added to the AI categorization sensory AI, cognitive AI, and physical AI. Sixth, based on the stage of AI development at a company level Gartner research team (2020) identified the following five types. The attentive AI is recognized by the company but not in a strategic way, and no pilot projects or experiments are taking place. Active AI is executed in proofs of concept and pilot projects. Operational AI is characterized with at least one AI project which has moved to production and a dedicated budget. Systematic AI is considered when AI-powered applications interact productively within the organization and across the business ecosystem. The highest level of AI is regarded as transformational AI.

Kendall (1999) categorized AI as a decision-support technology contrary to cooperation-enabling technologies and infrastructure-enabling technologies. A decision-supporting technology implies that it directly affects the capacity of an individual, organization, or team to create models efficiently, make more effective decisions, or develop alternatives and solutions.

McKinsey Global Institute (2017) provides an interesting categorization of companies depending on the level of AI implementation which is similar to the Roger's Innovation Adoption Curve. Front-runners, for example, tend to have a strong starting digital base, a higher propensity to invest in AI, and positive views for AI. Some AI innovators and creators which are kind of early adopters have big starting silo of data, computing power, and specialized talent. Other early adopters are innovative in how they deploy AI technologies. At the other end of the curve are the laggards that do not adopt AI technologies. It is expected that these type of companies will face a certain decline in their cash flow. This means that the companies will benefit disproportionately of AI adoption and implementation. One important driver of the profit pressure is the existence of strong competitive dynamics among firms.

There are various barriers which come across the development of artificial intelligence especially concerning knowledge acquisition, commonsense knowledge representation, and computing power. Since AI is still regarded as an emerging

information technology some typical barriers represent the uncertainty concerning the value of AI, the resistance or difficulty with use of AI, and the complexity of its implementation.

### **Societal Relevance of AI**

The intense penetration of AI in business raised the notion that "... AI deployment requires careful management to prevent unintentional but significant damage, not only to brand reputation but, more important, to workers, individuals, and society as a whole" (Burkhardt, Hohn and Wigley, 2019:2). The responsible development, dissemination, and use of AI knowledge for the benefit of society turned out to be a critical success factor both for the organizations and the institutions. Increasing societal focus on responsible use of AI systems inspired ethical debates about what is "right" and "wrong" when it comes to AI applications. Several perspectives within these ethical domains could be outlined.

The first domain deals with data issues such as data acquisition and data-set suitability. On the one hand, raising the amount of data used to train systems leads to more accurate and discerning predictions. On the other hand, protection of customers' individuality requires strict control on data acquisition, especially on third-party data and/or the repurposing of existing customer data. Sometimes despite the best of intentions of data scientists some data acquisition activities might be perceived by customers as invasion of privacy. Data-set suitability refers to the racial, gender, and other human biases, including time-selection bias as well. Data granularity is another important topic with this domain.

The second domain covers the whole AI development process, including data selection, feature selection, and model building and monitoring usually termed 'AI-output fairness'. The ethical debate is focused on machine learning techniques and algorithms since they have no awareness of the context in which their decisions will be applied. Some critical shifts regarding the transfer of responsibility are evident nowadays. Traditional machine learning techniques rely on a human to decide what aspects of the data are the most important to the model they are building. Though, some new methods rely on the machine to decide what is important in the data to drive the required outputs. Since this affects the design of the training and testing data, the most critical step is to establish definitions and metrics to evaluate fairness.

Regulatory compliance and engagement comprises the third domain. The main focus is placed on data-privacy protections. During the last few years the companies outside of regulated industries has been affected by various existing and emerging regulations, such as the General Data Protection Regulation (GDPR), the Children's Online Privacy Protection Act (COPPA), the California Consumer Privacy Act (CCPA). The changes in legislation require reassessment of organizations' policies



toward use of customer data and the right of customers to be evaluated by a human, the right to be forgotten, and automatic profiling. The successful implementation of these policies needs in turn a detailed definition of compliance metrics for AI initiatives. The compliance behavior of organizations could be regarded as lagging behavior or reactive approach. The proactive approach demands engagement of organizational management toward development of new regulations.

The fourth domain includes the AI model explainability<sup>1</sup> which exceeds the scope of organization demanding an explanation of AI model outputs to stakeholders. The success of the debates around the AI model explainability requires development of common language and shared understanding of the topics in order to solve the potential biases well before models are implemented.

AI models are usually evaluated on the following two criteria: interpretability and predictive performance (Burkhardt, Hohn and Wigley, 2019:7). It should be taken into account that models with more predictive power are often more opaque, e.g. neural networks. Linear regression, for example, is characterized with high interpretability but with extremely low predictive performance.

Recent work of World Economic Forum (2020) on AI ecosystem mapping provides useful insights on the interrelations between main AI domains and the corresponding responsibility and ethical issues. AI education and awareness are closely related to the issues of global risks, future of economic progress, public finance and social protection, education and skills, workforce and employment. AI ethics and values relate to global governance, international security, corporate governance, agile governance, global risks, education and skills, and healthcare delivery. Specific attention is given to AI safety, security and standards and their links to geo-economics, internet governance, future of media, entertainment and culture, leadership in the Fourth Industrial Revolution, global governance and international security.

Another aspect of societal relevance of AI reflects the uneven distribution of benefits of the AI-based economic activity at several levels – country, organization, workers/individuals (Van Dijk, 2012). AI may widen gaps between countries due to different AI adoption levels thus reinforcing the current digital divide. Large companies have a competitive advantage in adopting and absorbing AI ahead of industry peers. A widening gap may also appear at the level of individual workers because of the transfer from repetitive job activities to more creative work or activities which require more digital skills.

## **SOLUTIONS AND RECOMMENDATIONS**

As intelligence becomes a primary commodity, innovation, creativity, and ongoing learning are the currency of this new intelligence-based economy (Vandergriff, 2008:432). Machines are becoming more autonomous but they are, as Wallach and Allen (2009) said, ethically blind. Despite the positive impact on companies and the increasing societal relevance of AI, there has been a growing recognition of the potentially negative impact of artificial intelligence on society (WEForum, 2020). There is a certain consolidation of attitudes and opinions toward the following key principles of responsible AI:

respect for privacy, transparency, explainability, human control, and mitigating bias. The main challenge is to implement these principles into practice, ensure and encourage their use. Since these principles are general by nature, they should be operationalized properly. Moreover, their implementation will require a relevant accountability and auditing framework. Companies should protect their customers' individuality, their customers' freedom not to be predictable, their customers' privacy (Carmon, Schrift, Wertenbroch and Yang, 2019).

## **FUTURE RESEARCH DIRECTIONS**

Research in AI is both theoretical and applied, and transcends traditional disciplinary boundaries, bringing together experts from diverse fields of study (Cockburn et al., 2018). Despite the increasing adoption of AI tools and applications there are still some limitations which are pure technical, practical limitations and limitations in use. Future research directions will focus these general limitations. Regarding responsible AI, future research areas should cover the following topics: bias, transparency and fairness in AI algorithms; responsible AI operationalization, geopolitical impacts of AI.

## **CONCLUSION**

Alessandro Annoni<sup>2</sup> (2019) noted that "AI ... offers major opportunities to improve our lives but ethical and secure-by-design algorithms are crucial to building trust in this disruptive technology." The evolution of AI bursts with contradictions. AI possesses a huge potential to improve human lives, and, at the same time, it could widen the social and digital divides. In order to utilize its positive potential and to minimize the threats, it is critical to involve people (experts, scientists, policy makers, funders and investors, etc.) with various background and organizations from different

industries to build a solid background of common language and shared understanding of AI capabilities and risks to guide all stakeholders to positive impacts. A broader engagement of civil society on the values that need to be embedded in AI and the directions for future development are also needed. The key to success is to balance the transformational potential of AI with human safety and privacy.

## **REFERENCES**

Aayog, N. I. T. I. (2018). *National Strategy for Artificial Intelligence*. Discussion paper. [http://niti.gov.in/writereaddata/files/document\\_publication/NationalStrategy-for-AI-Discussion-Paper.pdf](http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf)

Andrews, W. (2020). *Build The AI Business Case*. Available at: <https://emtemp.gcom.cloud/ngw/globalassets/en/information-technology/documents/trends/ai-business-case-ebook.pdf>

Anyoha, R. (2020). *The History Of Artificial Intelligence - Science In The News*. Available at: <http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes From the AI Frontier: Modeling the Impact of AI on the World Economy*. McKinsey Global Institute. Available at: <https://mck.co/3aIkDZT>

Burkhardt, R., Hohn, N., & Wigley, C. (2019). *Leading Your Organization To Responsible AI*. Available at: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/leading-your-organization-to-responsible-ai>

Cam, A., Chui, M., & Hall, B. (2019). *Global AI Survey: AI Proves Its Worth, But Few Scale Impact*. Available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact>

Carmon, Z., Schrift, R., Wertenbroch, K., & Yang, H. (2019). *Designing AI Systems That Customers Won'T Hate*. MIT Sloan Management Review. Available at: <https://sloanreview.mit.edu/article/designing-ai-systems-that-customers-wont-hate/>

Chui, M., & Malhotra, S. (2018). *AI Adoption Advances, But Foundational Barriers Remain*. McKinsey Analytics. Available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain>

Chui, M., & Manyika, J. (2018). *The Real-World Potential And Limitations Of Artificial Intelligence*. Available at: <https://www.mckinsey.com/featured-insights/artificial-intelligence/the-real-world-potential-and-limitations-of-artificial-intelligence>

## **Artificial Intelligence**

Cockburn, I., Henderson, R., & Stern, S. (2018). *The Impact Of Artificial Intelligence On Innovation*. National Bureau of Economic Research. Working paper No. 2449. Available at: <https://www.nber.org/papers/w24449.pdf>

Copeland, J. (2019). *Artificial Intelligence: How Does AI Work?* Independently Published.

Ein-Dor, P. (1999). Artificial Intelligence: A Short History and the Next 40 Years. In K. Kendall (Ed.), *Emerging Information Technologies*. Sage Publications.

Ein-Dor, P. (2011). Taxonomies of Knowledge. In *Encyclopedia of Knowledge Management* (2nd ed., pp. 1490–1499). IGI Global.

Hager, G., Drobnis, A., Fang, F., Ghani, R., Greenwald, A., Lyons, T., Parkes, D., Schultz, J., Saria, S., Smith, S., & Tambe, M. (2017). *Artificial Intelligence for Social Good*. Computing Community Consortium. <https://cra.org/ccc/wp-content/uploads/sites/2/2016/04/AI-for-Social-Good-Workshop-Report.pdf>

Kendall, K. (1999). *Emerging Information Technologies*. Sage Publications.

Miguel Furtado Cardoso Lopes, G. (2009). Gordon Pask: Exchanges between cybernetics and architecture and the envisioning of the IE. *Kybernetes*, 38(7/8), 1317–1331.

Nilsson, N. (2010). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge University Press.

Olley, D. (2020). *AI Report | Research Intelligence | Elsevier*. Available at: <https://www.elsevier.com/research-intelligence/resource-library/ai-report>

Panetta, K. (2020). *The CIO'S Guide To Artificial Intelligence*. <https://www.gartner.com/smarterwithgartner/the-cios-guide-to-artificial-intelligence/>

Russell, S., & Norvig, P. (2010). *Artificial Intelligence*. Prentice Hall.

Schwab, K. (2016). *The 4th Industrial Revolution*. New York, NY: World Economic Forum.

Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J., Lyons, T., Etchemendy, J., Grosz, B., & Bauer, Z. (2018). *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University.

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). *Artificial Intelligence And Life In 2030. One Hundred Year Study On Artificial Intelligence. Report of the 2015-2016 Study Panel*. Stanford University. Available at <http://ai100.stanford.edu/2016-report>

van Dijk, J. A. G. M. (2012). The Evolution of the Digital Divide - The Digital Divide Turns to Inequality of Skills and Usage. In J. Bus, M. Crompton, M. Hildebrandt, & G. Metakides (Eds.), *Digital Enlightenment Yearbook 2012* (pp. 57-78). Amsterdam: IOS Press.

Vandergriff, L. (2008). Welcome to the Intelligence Age: An examination of intelligence as a complex venture emergent behavior. *Vine*, 38(4), 432–444.

Wallach, W., & Allen, C. (2009). *Moral Machines*. Oxford University Press.

World Economic Forum. (n.d.). *Artificial Intelligence and Robots*. <https://toplink.weforum.org/knowledge/insight/a1Gb0000000pTDREA2/explore/summary>

## **ADDITIONAL READING**

Ai100.stanford.edu. (2020). *2016 Report | One Hundred Year Study On Artificial Intelligence (AI100)*. Available at: <https://ai100.stanford.edu/2016-report>

## **KEY TERMS AND DEFINITIONS**

**Deep Learning:** Deep learning is a type of machine learning that can process a wider range of data resources, requires less data preprocessing by humans, and can often produce more accurate results than traditional machine-learning approaches. In deep learning, interconnected layers of software-based calculators known as “neurons” form a neural network. The network can ingest vast amounts of input data and process them through multiple layers that learn increasingly complex features of the data at each layer. The network can then make a determination about the data, learn if its determination is correct, and use what it has learned to make determinations about new data. For example, once it learns what an object looks like, it can recognize the object in a new image (Source: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai>).

**ICTAIs:** International Centres for Transformational AI. The priority domains of ICTAIs include agriculture, health, education, smart cities and infrastructure, smart mobility and transportation (Source: [http://niti.gov.in/writereaddata/files/document\\_publication/NationalStrategy-for-AI-Discussion-Paper.pdf](http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf)).

**LIME (Local Interpretable Model-Agnostic Explanations):** LIME is short for Local Interpretable Model-Agnostic Explanations. Each part of the name reflects something that characterizes the model. Local refers to local fidelity - i.e., we want the explanation to really reflect the behaviour of the classifier “around” the instance being predicted. This explanation is useless unless it is interpretable - that is, unless a human can make sense of it. LIME is able to explain any model without needing to ‘peak’ into it, so it is model-agnostic. LIME is a technique to explain the predictions of any machine learning classifier, and evaluate its usefulness in various tasks related to trust (Source: <https://bit.ly/36SmFbz>).

**Machine Learning:** Most recent advances in AI have been achieved by applying machine learning to very large data sets. Machine-learning algorithms detect patterns and learn how to make predictions and recommendations by processing data and experiences, rather than by receiving explicit programming instruction. The algorithms also adapt in response to new data and experiences to improve efficacy over time (Source: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai>).

**SHAP (SHapley Additive exPlanations):** SHAP is a method to explain individual predictions. It is based on the game theoretically optimal Shapley Values. The goal of SHAP is to explain the prediction of an instance  $x$  by computing the contribution of each feature to the prediction. The SHAP explanation method computes Shapley values from coalitional game theory. The feature values of a data instance act as players in a coalition (Sources: Lundberg, Scott M., and Su-In Lee. (2017) A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems; Molnar, C. (2019). Interpretable machine learning. A Guide for Making Black Box Models Explainable. <https://christophm.github.io/interpretable-ml-book/>).


## ENDNOTES

- <sup>1</sup> The field of so called “explainable AI” (sometimes referred to as XAI) is quite new but it is rapidly maturing. Various tools are developed such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations), and activation atlases.
- <sup>2</sup> Head of Digital Economy Unit, Joint Research Centre, European Commission, [https://digitalearth2019.eu/alessandro\\_annoni/](https://digitalearth2019.eu/alessandro_annoni/); <https://www.elsevier.com/?a=823654>

## Chapter 2

# The Soul of Artificial Intelligence and Races' Separation of AI and Homo

**Rinat Galiautdinov**

 <https://orcid.org/0000-0001-9557-5250>  
Independent Researcher, Italy

### ABSTRACT

*Artificial intelligence breaks into our lives. However, some questions are already being raised, and increasingly, these issues affect aspects of morality and ethics. Is it possible to scoff at thinking AI? When will it be invented? What prevents us from writing laws of robotics right now, putting morality into them? What surprises does machine learning bring us now? Can machine learning be fooled, and how difficult is it? But even greater questions arise in the context of the separation of races of AI and humans. Is AI racism our future? This chapter explores these questions.*

### INTRODUCTION

New technologies are changing our daily lives and raising ethical issues that did not exist before. The changes in the life of mankind that artificial intelligence can bring and is already bringing are difficult to compare with what appeared earlier. Humanity can get rid of most of the well-known professions, and potentially create a new form of life. Julia Bossmann, president of Foresight Institute, based in Palo Alto, who promotes transformative technology, tried to describe the ethical issues that may arise during this process.

DOI: 10.4018/978-1-7998-4285-9.ch002

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



Logistics optimization, fraud detection, research and translation: smart computer systems are changing our lives for the better. The more capable they become, the more efficiently our world works, and therefore richer.

Technology giants like Alphabet (Google), Amazon, Facebook, IBM and Microsoft, as well as individuals like Stephen Hawking and Elon Musk, believe that now is the right time to discuss the almost limitless landscape of artificial intelligence. In many cases, this is as much a new frontier for ethics and risk assessment as it is for new technologies. So what are the problems and conversations that keep AI experts awake?

And let's also consider another problem related to the digital avatar of a human. Here in this context under the digital avatar of a human the author means that at some point of time we will be able to digitize the human brain including all the human's knowledge and thoughts, so that this digital representation of a human (having the ability to think and obviously possessing the virtual nervous system) which implementation was described in the number of the researches of the author (Galiautdinov Rinat, 2020; Galiautdinov Rinat & Mkrttchian Vardan, 2019 A; Galiautdinov Rinat & Mkrttchian Vardan, 2019 B).

The digital avatar of a person could continue living even after a dead of the person and could act either as a memory keeper of the person or as a virtual member of society, who possess the knowledge of the previous epochs but also continues his/her education and progresses in the other spheres.

So from this prospective the future society could look as a combination of live people and virtual avatars which creates the constant interactions and competitions between them. At some point we can even consider the separation of the human race into 2 major groups: live people and virtual avatars.

Virtual avatars can and possess the pretty much of what possess the robots build on the basis of Artificial Intelligence however the robots do not possess the human personality and real life experience of a natural creature, such as homo.

The technical implementation contains lots of the challenges and although it's possible to create the avatar of a simple natural creature, such as Aplysia (the mollusk), it's still impossible to create something more complex, for example the avatar of a mouse or human being.

But this is a question of time and sooner or later it will be done.

Which creates the number of the ethical issues:

- Will such the digital avatar possessing AI be responsible?
- Will a human be responsible to destroying of the digital Avatar?
- Will digital avatar change his/her way of thinking over the hundred/thousands of years?

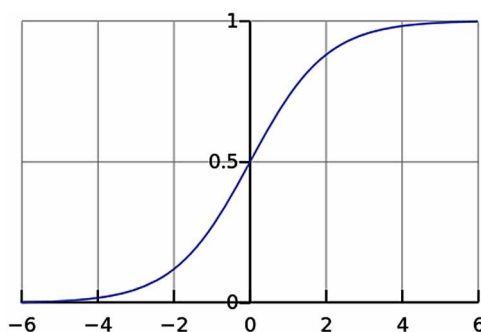
- Will digital avatar's evolution road parted from the evolution and mental and moral evolution of humans?
- Will digital avatar start the war against humans?
- Will digital avatar accept the idea that a digital avatar is more important than a human?
- Will digital avatars decide to separate from humans? Will they create their own country? Will they hire humans for doing only some low quality work related to maintenance of the computers and digital devices where digital avatars actually live?
- Will they decide to get rid of humans once they create robotic body?

But first let's consider the current state of the AI. There are two different things: Strong and Weak AI.

Strong AI (true, general, real) is a hypothetical machine that can think and be aware of itself, solve not only highly specialized tasks, but also learn something new. Weak AI (narrow, shallow) - these are already existing programs for solving quite specific tasks, such as image recognition, auto driving, playing Go, etc. In order not to get confused and not to mislead anyone, we prefer to call the Weak AI "machine learning" (machine learning). About Strong AI, it is still unknown whether it will ever be invented at all. On the one hand, until now, technologies have developed with acceleration, and if this goes on, then there are five years left. On the other hand, few processes in nature actually proceed exponentially. Much more often, after all, we see a logistic curve.

While we are somewhere on the left of the chart, it seems to us that this is an exponent. For example, until recently, the world's population has grown with such acceleration. But at some point "saturation" occurs, and growth slows down. When experts are questioned, it turns out that on average, wait another 45 years.

*Figure 1.*



Curiously, North American scientists believe that AI will surpass humans in 74 years, and Asian scientists in just 30. Perhaps in Asia they know something ...

These same scientists predicted that a machine would translate better than a person by 2024, write school essays by 2026, drive trucks by 2027, play Go by 2027 too. Go has already missed, because this moment came in 2017, only 2 years after the forecast.

## **PROBLEMS**

Let's consider the most obvious problems we will face with along with the development of the AI.

### **Unemployment**

What will happen to the extinction of professions? The hierarchy of labor is mainly associated with automation. Inventing new ways to automate work, we can give a place to new, more complex professions, moving from physical work that dominated the pre-industrial world to the cognitive work that is characteristic of the strategic and administrative work of our globalized society.

Take a look at trucking: in this area, millions of people work only in the United States. What will happen to them if the unmanned trucks promised by Elon Musk become widely available within ten years? On the other hand, if we take into account the reduced risk of car crashes, unmanned trucks look like an ethical choice. The same scenario can be applied to office workers, as to the majority of the workforce in developed countries.

And here we come to the question of how we will spend our time. Most people rely on selling their time to earn income and support themselves and their family. We can only hope that this opportunity will allow people to find meaning in unearned activities, such as caring for the family, interacting with the community and learning new ways to contribute to human society. If we make a successful transition, one day we can look around and think how barbaric it was that people had to sell most of their conscious time just to continue living.

### **Inequality**

How do we distribute the good produced by machines?

Our economic system is based on compensation for contributions to the economy, often measured by hourly wages. Most companies are still dependent on hourly labor when it comes to products and services. But using artificial intelligence, a company

can significantly reduce its dependence on human labor, which means fewer people will receive income. As a result, persons owning companies where the work is done by AI will receive all the money.

We already see a widening wealth gap, where startup founders get most of the economic surplus they create. In 2014, the three largest Detroit companies generated almost the same income as the three largest companies in Silicon Valley ... only in the Valley, they employed 10 times fewer employees.

If we imagine a post-working society, how will we structure an honest post-labor economy?

## **Humanity**

How do cars affect our behavior and interaction?

Artificial intelligence bots are getting better at modeling human conversation and relationships. In 2015, a bot named Evgeny Gustman (which was created by a Ukrainian and a Russian - ed.) Won the Turing contest for the first time in history. In it, people through text messages communicated with an unknown entity, and then tried to guess if they were talking with a person or a machine. Eugene Gustman circled around half the finger of his interlocutor, making them think that they were talking to the person.

This stage is only the beginning of an era where we will often interact with machines as if they were people. While people are limited in attention and kindness, which they can spend on another person, bots can spend almost unlimited resources on building relationships.

Despite the fact that many do not know this, we have already become victims of how machines can make the reward center in the human brain work. Just take a look at clickbait titles and video games. These headers are often optimized through A / B testing, a rudimentary form of algorithmic optimization used to make our content catch our eye better. This and other methods are used to make many video and mobile games addictive. Technological dependence is a new facet of human dependence.

On the other hand, perhaps we can come up with another way of using software that can already effectively direct human attention and cause certain actions. Its proper use can lead to the possibility of directing society to more useful behavior. However, in the wrong hands, this can be harmful.

## **Artificial Dullness**

How can we protect ourselves from mistakes?

Intelligence will come about through training, regardless of whether you are a person or a machine. Systems usually go through the training phase, where they

“learn” to find the correct patterns and act according to the data entered. Once the system is fully trained, it can go into the test phase, where they will be given various examples and watch how it copes with them.

Obviously, the training phase cannot cover all possible examples that the system may encounter in the real world. Such systems can be fooled in a way that people cannot be fooled. For example, random sets of points can cause machines to “see” things that aren’t there. If we rely on AI to bring us into the world of new work, safety and efficiency, we need to make sure that the machines behave as planned and that people cannot take control of them in order to use them for personal purposes.

## **Racist Robots**

How do we get rid of AI bias?

Despite the fact that the capabilities of artificial intelligence in the speed and amount of data processing are significantly superior to human, it can not always be trusted as honest and neutral. Google and its parent company Alphabet are leaders in the field of artificial intelligence. This can be seen on Google Photos, where AI is used to identify people, objects, and scenes. But something may go wrong. As in the case when the camera erroneously detected blinking in people of Asian appearance, or the software used to predict future criminals, which was biased against blacks.

We must not forget that AI-based systems are created by people who may be biased. But again, if they are used correctly or if they are used by people striving for social progress, they can become a catalyst for positive changes.

## **Security**

How to keep AI safe from opponents?

The more powerful the technology becomes, the more likely it is to use it for low purposes. This applies to both robots designed to replace human soldiers or autonomous weapons, and AI systems that can do harm if they are used maliciously. As the battles with their use will unfold not only in the fields, cybersecurity will become increasingly important. In the end, we are dealing with a system that is faster and more capable of us at times.

## **Singularity**

How can we control a complex smart system?

The reason people are at the top of the food chain is not because we have sharp teeth or strong muscles. The dominance of mankind is built almost entirely on our ingenuity and intelligence. We can get the best from larger, faster and stronger animals,

because we can create tools to control them: physical, like cages and weapons, and cognitive, like training.

This poses a serious question about artificial intelligence: will it one day be able to get the same advantage over us? We cannot rely solely on pulling the plug out of the socket, because a sufficiently advanced machine will be able to anticipate this step and protect itself. This is what is called “singularity”: the moment when human beings cease to be the most intelligent on Earth.

## **The Rights of Robots: How Do We Define Human Attitudes Toward AI?**

While neurophysiologists are working to unravel the secrets of conscious experience, we better understand the basic mechanisms of reward and rejection. Even the simplest animals have these mechanisms. In a sense, we are creating similar mechanisms in artificial intelligence systems. For example, reinforcement training is like training a dog: progress in completing a task is reinforced by a virtual reward.

Right now, such systems are quite superficial, but they are becoming more complex and lively. Can we assume that a system suffers if its reward function is given a negative input? Moreover, the so-called genetic algorithms work by creating multiple instances of the system at the same time. Of these, only the most successful “survive” in order to unite and form a new generation of specimens. This has been going on for generations and is a means of improving the system. Unsuccessful instances are deleted. At what point can we view genetic algorithms as a form of massacre?

As soon as we begin to consider machines as entities that can perceive, feel and act, we will need to think about changing their legal status. Should they be treated like animals with comparable intelligence? Will we take into account the suffering of the “sentient” machine?

Some ethical issues relate to mitigating efforts, some relate to the risk of negative consequences. As we look at these risks, we must also keep in mind that overall this technological advancement means a better life for everyone. Artificial intelligence has enormous potential, and its responsible use depends on us.

## **Strong AI Raises a Lot of Ethical Issues**

Although Strong AI will wait a long time, but we know for sure that there will be enough ethical problems. The first class of problems is that we can offend AI. For example:

- Is it ethical to torture AI if it can feel pain?

- Is it normal to leave AI without communication for a long time if it is able to feel loneliness?
- Can you use it as a pet? What about a slave? And who will control this and how, because this is a program that works “lives” in your “smart-phone”?

Now no one will be outraged if you offend your voice assistant, but if you mistreat the dog, you will be condemned. And this is not because she is of flesh and blood, but because she feels and experiences a bad attitude, as it will be with Strong AI.

The second class of ethical issues - AI can offend us. Hundreds of such examples can be found in films and books. How to explain AI, what do we want from it? People for AI are like ants for workers building a dam: for the sake of a great goal, you can crush a couple.

Science fiction plays a trick on us. We are used to thinking that Skynet and the Terminators are not there, and they will not be soon, but for now you can relax. The AI in films is often malicious, and we hope that this will not happen in life: after all, we were warned, and we are not as stupid as the heroes of the films. Moreover, in thoughts about the future, we forget to think well about the present.

Machine learning allows you to solve a practical problem without explicit programming, but through training on precedents. Since we are teaching a machine to solve a specific problem, the resulting mathematical model (the so-called algorithm) cannot suddenly want to enslave / save humanity. Do it normally - it will be normal. What could go wrong?

## **Bad Intentions**

First, the task itself may not be ethical enough. For example, if we teach machine drones to kill people using machine learning. Just recently, a little scandal broke out about this. Google is developing the software used for the Project Maven drone management pilot project. Presumably in the future this could lead to the creation of a fully autonomous weapon.

So, at least 12 Google employees quit in protest, another 4,000 signed a petition asking them to abandon the contract with the military. More than 1000 prominent scientists in the field of AI, ethics and information technology wrote an open letter asking Google to stop working on the project and support the international treaty banning autonomous weapons.

But even if the authors of the machine learning algorithm do not want to kill people and do harm, they nevertheless often still want to make a profit. In other words, not all algorithms work for the benefit of society, many work for the benefit of creators. This can often be observed in the field of medicine - it's more important

not to cure, but to recommend more treatment. In general, if machine learning recommends something paid, it is very likely that the algorithm is greedy.

Well, and sometimes society itself is not interested in the resulting algorithm being a model of morality. For example, there is a trade-off between vehicle speed and road deaths. We could greatly reduce mortality if we limited the speed to 15 mph, but then life in big cities would be difficult.

## **Ethics Is Only One of the Parameters of the System**

Imagine, we ask the algorithm to make up the country's budget in order to "maximize GDP / labor productivity / life expectancy". There are no ethical limitations and goals in the formulation of this task. Why allocate money for orphanages / hospices / environmental protection, because it will not increase GDP (at least directly)? And it's good if we only entrust the budget to the algorithm, because in a broader statement of the problem it turns out that an unemployed population is "more profitable" to kill immediately in order to increase labor productivity. It turns out that ethical issues should be among the goals of the system initially.

## **Ethics Is Hard to Describe Formally**

There is one problem with ethics - it is difficult to formalize. Different countries have different ethics. It changes over time. For example, on issues such as gays rights and interracial / inter-caste marriages, opinions can change significantly over decades. Ethics may depend on the political climate.

For example, in China, monitoring the movement of citizens using surveillance cameras and face recognition is considered the norm. In other countries, the attitude to this issue may be different and depend on the situation.

## **Machine Learning Affects People**

Imagine a machine-learning-based system that advises you which movie to watch. Based on your ratings for other films, and by comparing your tastes with those of other users, the system can quite reliably recommend a movie that you really like.

But at the same time, the system will change your tastes over time and make them more narrow. Without a system, from time to time you would watch bad films and films of unusual genres. And so that no movie - to the point. As a result, we cease to be "film experts", and become only consumers of what they give. It is also interesting that we do not even notice how the algorithms manipulate us.

If you say that such an effect of algorithms on people is even good, then here is another example. China is preparing to launch the Social Rating System - a system



for evaluating individuals or organizations by various parameters, the values of which are obtained using mass surveillance tools and using big data analysis technology. If a person buys diapers - that's good, the rating is growing. If spending money on video games is bad, the rating drops. If communicating with a person with a low rating, then also falls.

As a result, it turns out that thanks to the System, citizens consciously or subconsciously begin to behave differently. Communicate less with unreliable citizens, buy more diapers, etc.

## **Algorithmic System Error**

Besides the fact that sometimes we ourselves don't know what we want from the algorithm, there is also a whole bunch of technical limitations.

The algorithm absorbs the imperfection of the world. If we use data from a company with racist politicians as a training sample for the hiring algorithm, then the algorithm will also be racist.

Microsoft once taught a chatbot to chat on Twitter's. It had to be turned off in less than a day, because the bot quickly mastered curses and racist remarks.

In addition, the learning algorithm cannot take into account some unformalized parameters. For example, when calculating the recommendation to the defendant - to admit or not to admit guilt on the basis of the evidence gathered, it is difficult for the algorithm to take into account how impressed the admission will make to the judge, because the impression and emotions are not recorded anywhere.

## **False Correlations and Feedback Loops**

A false correlation is when it seems that the more firefighters in a city, the more often fires. Or when it is obvious that the fewer pirates on Earth, the warmer the climate on the planet.

So - people suspect that pirates and the climate are not directly connected, and it is not so simple with firefighters, and the machine learning model simply memorizes and generalizes. Well-known example. The program, which arranged patients in turn according to the urgency of relief, concluded that asthmatics with pneumonia need less help than just people with pneumonia without asthma. The program looked at the statistics and came to the conclusion that asthmatics do not die - why do they need priority? And they do not really die because such patients immediately receive the best care in medical institutions due to a very high risk (Kuperman et al., 2006).

Worse than false correlations are only feedback loops. A California crime prevention program suggested sending more cops to black neighborhoods based on crime rates (number of reported crimes). And the more police cars in the field of

visibility, the more often residents report crimes (just have someone to report). As a result, crime is only increasing - that means more police officers must be sent, etc.

In other words, if racial discrimination is a factor of arrest, then feedback loops can strengthen and perpetuate racial discrimination in police activities.

## **Who to Blame**

In 2016, the Big Data Working Group under the Obama Administration issued a report warning of “the possible coding of discrimination in making automated decisions” and postulating the “principle of equal opportunity”.

But to say something is easy, but what to do?

First, machine learning math models are hard to test and tweak. For example, the Google Photo app recognized people with black skin like gorillas. And what to do? If we read ordinary programs step by step and learned how to test them, then in the case of machine learning, everything depends on the size of the control sample, and it cannot be infinite. For three years, Google could not come up with anything better than to turn off the recognition of gorillas, chimpanzees and monkeys at all, so as to prevent a repeat of the error.

Secondly, it is difficult for us to understand and explain machine learning solutions. For example, a neural network somehow placed weight coefficients within itself to get the correct answers. And why do they turn out just like that and what to do to change the answer?

A 2015 study found that women are much less likely than men to see high-paying job postings advertised by Google AdSense. Amazon's same-day delivery service was regularly unavailable in the black quarters. In both cases, company representatives found it difficult to explain such solutions to the algorithms.

## **It Remains to Make Laws and Rely on Machine Learning**

It turns out that there is no one to blame, it remains to pass laws and postulate the “ethical laws of robotics.” Some European country recently issued such a set of rules for unmanned vehicles. Among other things, it says:

- Human safety is the highest priority compared to damage to animals or property.
- In the event of an imminent accident, there should be no discrimination, for no reason it is unacceptable to distinguish between people.

But what is especially important in our context:

Automatic driving systems become an ethical imperative if systems cause fewer crashes than human drivers. Obviously, we will increasingly rely on machine learning - simply because it will generally do better than humans.

## **Machine Learning Can Be Poisoned**

And here we come to no lesser misfortune than the bias of the algorithms - they can be manipulated.

Machine Learning Poisoning (ML poisoning) means that if someone takes part in the training of the model, then he can influence the decisions made by the model.

For example, in a computer virus analysis laboratory, a model processes an average of a million new samples every day (clean and malicious files). The threat landscape is constantly changing, therefore changes in the model in the form of anti-virus database updates are delivered to the anti-virus products on the user side (Wendell W., 2010; Prakken H., 2017).

So, an attacker can constantly generate malicious files very similar to some clean one and send them to the laboratory. The border between clean and malicious files will gradually be erased, the model will “degrade”. And in the end, the model can recognize the original clean file as malicious - it will result in a false positive.

And vice versa, if you “spam” a self-learning spam filter of a ton of clean generated emails, you will eventually be able to create spam that passes through the filter (Smoliar et al. 1994).

Therefore, a multi-level approach to protection is necessary, we do not rely only on machine learning.

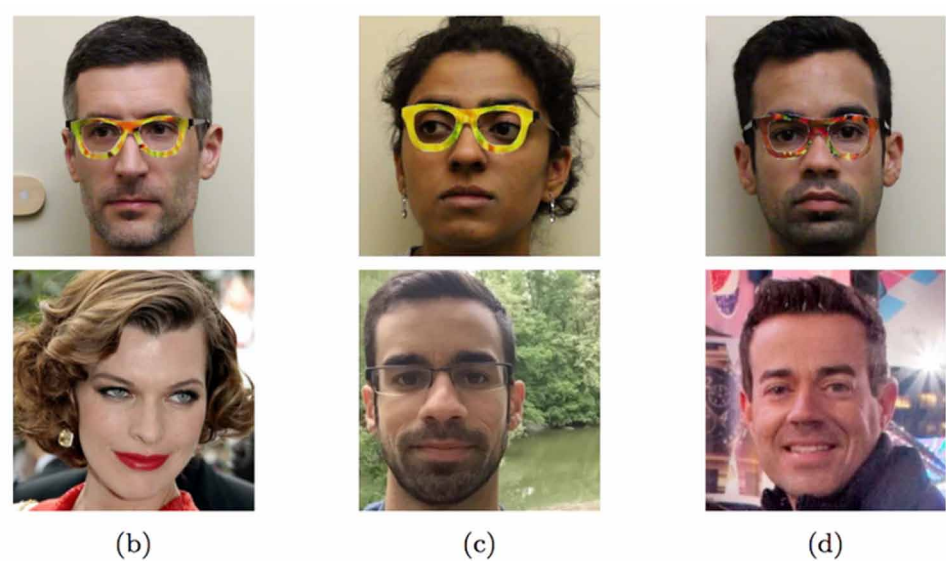
Another example, while fictional. In the face recognition system, you can add specially generated faces, so that in the end the system begins to confuse you with someone else. Do not think that this is impossible, take a look at the picture from the next section.

## **Machine Learning Hacking**

Poisoning is an effect on the learning process. But it is not necessary to participate in training in order to get a benefit - you can also deceive a ready-made model if you know how it works.

Figure 2 Illustrates that when wearing specially colored glasses, the researchers posed as other famous people. This example with faces has not yet been encountered in the wild - precisely because no one has yet entrusted the machine with making important decisions based on face recognition. Without human control, it will be exactly as in the picture.

*Figure 2.*



Even where, it would seem, there is nothing complicated, it is easy to deceive a car in an unknown way to the uninitiated. (Knight W., 2017; Neumann et al., 2008)

Figure 3 Illustrates that the first three characters are recognized as “Speed Limit 45” and the last as STOP. Moreover, in order for the machine learning model to recognize surrender, it is not necessary to make significant changes, it is enough minimal edits that are invisible to a person.

*Figure 3.*



Figure 4.

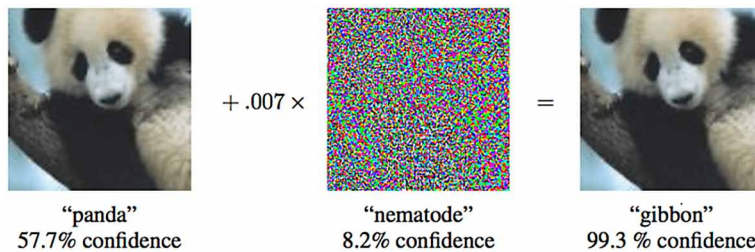


Figure 4 Illustrates that If you add minimal special noise to the panda on the left, then machine learning will be sure that it is a gibbon. While a person is smarter than most algorithms, he can cheat them. Similarly, it will be possible to “hack” the Chinese Social Rating System and become the most respected person in China.

## Ethical Challenges of Digital Avatar

However the created digital avatar creates different kinds of the ethical challenges. A live digital avatar will possess the bunch of different kind of sensors and consequently will constantly receive and process the information and at some point it will start changing itself: the virtual nervous system of avatar will create the new synaptic contacts, will erase the old ones, will enhance some other synaptic contacts and will reduce the “power” of the other synaptic contacts. So after a while we might have some avatar who has nothing to do with his original model.

In the case if we want to avoid such the avatar evolution, we would need to limit the ability of the nervous system of the Avatar. However in this case the avatar will have limited abilities with quite poor Artificial Intelligence which eventually will become the barrier for the AI.

## SUMMARY

Let’s summarize what was managed to discuss in the article.

There is no strong AI yet.

We are relaxed.

Machine learning will reduce the number of victims in critical areas.

We will rely on machine learning more and more.

We will have good intentions.

We will even lay ethics in system design.

But ethics is hard formalized and different in different countries.

Machine learning is full of bias for various reasons.

We cannot always explain the solutions of machine learning algorithms.

Machine learning can be poisoned.

And even “hack”.

An attacker can gain an advantage over other people in this way.

Machine learning has an impact on our lives. Digital Avatar will continue its evolution.

Digital Avatar will eventually separate itself from a human race.

Digital Avatar will eventually use people, as more primitive creatures, for their own purposes (generally related to maintenance, etc.)

And all this is the near future.

## **REFERENCES**

Knight, W. (2017). *Boston may be famous for bad drivers, but it's the testing ground for a smarter self-driving car*. MIT Technology Review.

Kuperman, G. J., Reichley, R. M., & Bailey, T. C. (2006). Using Commercial Knowledge Bases for Clinical Decision Support: Opportunities, Hurdles, and Recommendations. *Journal of the American Medical Informatics Association*, 13(4), 369–371. doi:10.1197/jamia.M2055 PMID:16622160

Neumann, B., & Moller, R. (2008). On scene interpretation with description logics. *Image and Vision Computing*, 26(1), 82–101. doi:10.1016/j.imavis.2007.08.013

Prakken, H. (2017). On the problem of making autonomous vehicles conform to traffic law. *Artificial Intelligence and Law*, 25(3), 341–363. doi:10.1007/10506-017-9210-0

Rinat, G. (2020). Brain machine interface: The accurate interpretation of neurotransmitters' signals targeting the muscles. *International Journal of Applied Research in Bioinformatics*, 0102. Advance online publication. doi:10.4018/IJARB.2020

Rinat & Vardan. (2019a). Math model of neuron and nervous system research, based on AI constructor creating virtual neural circuits: Theoretical and Methodological Aspects. In V. Mkrttchian, E. Aleshina, & L. Gamidullaeva (Eds.), *Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms* (pp. 320–344). IGI Global. doi:10.4018/978-1-7998-1581-5.ch015

Rinat & Vardan. (2019b). Brain machine interface – for Avatar Control & Estimation in Educational purposes Based on Neural AI plugs: Theoretical and Methodological Aspects. In V. Mkrttchian, E. Aleshina, & L. Gamidullaeva (Eds.), *Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms* (pp. 345–360). IGI Global. doi:10.4018/978-1-7998-1581-5.ch016

Smoliar, S. W., & HongJiang Zhang. (1994). Content based video indexing and retrieval. *IEEE MultiMedia*, 1(2), 62–72. doi:10.1109/93.311653

Vardan, M., Leyla, G., & Rinat, G. (2019). Design of Nano-scale Electrodes and Development of Avatar-Based Control System for Energy-Efficient Power Engineering: Application of an Internet of Things and People (IOTAP) Research Center. *International Journal of Applied Nanotechnology Research*. Advance online publication. doi:10.4018/IJANR.201901010

Wendell, W. (2010). *Moral Machines*. Oxford University Press.

# Chapter 3

## Practical Issues in Human and Artificial Intelligence Interaction

**Arthur Kordon**  
*Kordon Consulting LLC, USA*

### **ABSTRACT**

*The chapter will focus on some practical issues in human and AI interaction based on the experience of applying AI in several large corporations. The following issues will be discussed: weaknesses of human intelligence, weaknesses of AI, benefits of human intelligence from AI, negative effects of AI on human intelligence, resistance of human intelligence toward AI, and how to improve the interaction between human and artificial intelligence. The discussed issues will be illustrated with examples from real-world applications.*

### **INTRODUCTION**

The fast invasion of Artificial Intelligence (AI) in industry caught the businesses unprepared. While the current focus of interest of industry is on understanding the various AI-related technologies and their value creation capabilities, the long-term effects of the interaction between human intelligence and the applied AI-based systems gradually grabs the attention of researchers and practitioners. Analyzing the different aspects of this complex phenomenon is of critical importance for the future productivity of the business applications of AI. For example Google was forced to admit in a note to investors that products and services “that incorporate or utilize artificial intelligence and machine learning, can raise new or exacerbate existing

DOI: 10.4018/978-1-7998-4285-9.ch003

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



ethical, technological, legal, and other challenges (Markus and Davis, 2019)”. There is a growing concern about human and artificial intelligence interaction in one of the key AI application – self-driving cars. It is not that it is impossible in principle to build a physical device based on AI that can drive in the snow or manage to be ethical; it is that we cannot get there with big data alone (Markus and Davis, 2019). Agreements, rules, and standards are beginning to emerge for issues such as user privacy, data exchange, and avoiding racial bias. Governments and corporations are working hard to sort out the rules for self-driving cars. There is a consensus that AI decisions must be explainable if AI systems are to be trusted, and that consensus is already partially implemented in the European Union’s General Data Protection Regulation (GDPR) legislation (Russel, 2019). There is a growing perception that our brains will be challenged to coevolve with our AI-rich devices in order to keep pace with exponentially accelerating intelligent machines (Coleman, 2019).

The author believes that this complex and important problem AI-based business applications needs a good understanding of the key features of AI and their relevance to human intelligence. A suggested starting point could be the 10 themes for responsible AI, defined in (Clarke, 2019):

1. Assess Positive and Negative Impacts and Implications
2. Complement Humans
3. Ensure Human Control
4. Ensure Human Safety and Wellbeing
5. Ensure Consistency with Human Values and Human Rights
6. Deliver Transparency and Auditability
7. Embed Quality Assurance
8. Exhibit Robustness and Resilience
9. Ensure Accountability for Obligations
10. Enforce, and Accept Enforcement of Liabilities and Sanctions

Based on these 10 themes Clarke formulates 50 principles for AI, which can be a good checklist to see what is relevant for the purpose when an AI-system is being used as it integrates all relevant aspects and stresses to deliver transparency, and auditability (Clarke, 2019). Taking into account these principles, the chapter will focus on some practical issues in human and AI interaction based on the experience of applying AI in several large corporations. The following issues are discussed:

- What are the key weaknesses of human intelligence that reduce the efficiency of business applications?
- What are the key benefits of AI that could improve human intelligence productivity?

- What are the key weaknesses and negative effects of AI that could reduce the efficiency of AI in real-world applications?
- Practical solutions how to improve the interaction between human and artificial intelligence.

The structure of the chapter follows these topics.

## **KEY WEAKNESSES OF HUMAN INTELLIGENCE**

It is not a secret that there is no universally accepted definition of human intelligence. Just as it is not fully comprehend how AI works, it is also not fully understood how human brains function, nor do we have a definitive grasp of what consciousness is, nor who, or possibly what, is conscious (Coleman, 2019).

The Wikipedia definition of human intelligence is ([https://en.wikipedia.org/wiki/Human\\_intelligence](https://en.wikipedia.org/wiki/Human_intelligence)):

*Human intelligence is the intellectual prowess of humans, which is marked by complex cognitive feats and high levels of motivation and self-awareness. Through their intelligence, humans possess the cognitive abilities to learn, form concepts, understand, apply logic, and reason, including the capacities to recognize patterns, comprehend ideas, plan, solve problems, make decisions, retain information, and use language to communicate.*

These extraordinary abilities are not unlimited, however. The following well-known flaws of human intelligence with direct impact in the interaction will be considered:

### **Cognitive Biases**

A cognitive bias is a systematic pattern of deviation from norm or rationality in judgment. It may sometimes lead to perceptual distortion, inaccurate judgment, illogical interpretation, or what is broadly called irrationality. Individuals create their own “subjective social reality” from their perception of the world. Individual’s construction of social reality, not the objective input from it, may dictate their behavior in the social world. Thus, cognitive biases may lead to distortion, inaccurate judgment, illogical interpretation, or what is broadly called irrationality.

The most important cognitive biases, related to AI and defined in Wikipedia, are listed below ([https://en.wikipedia.org/wiki/List\\_of\\_cognitive\\_biases](https://en.wikipedia.org/wiki/List_of_cognitive_biases)):

- *Dunning–Kruger effect*. The tendency for unskilled individuals to overestimate their own ability and the tendency for experts to underestimate their own ability.
- *Bandwagon effect*. The tendency to do (or believe) things because many other people do (or believe) the same.
- *Recency effect*. The tendency to weigh the latest information more heavily than older data.
- *Confirmation bias*. The tendency to search for, interpret, focus on, and remember information in a way that confirms one’s preconceptions.
- *Ostrich effect*. The tendency to ignore an obvious (negative) situation.
- *Ambiguity effect*. The tendency to avoid options for which missing information makes the probability seem “unknown.”
- *Clustering illusion*. The tendency to see patterns in random events (data).
- *Status quo bias*. The tendency to like things to stay relatively the same.
- *Weber–Fechner law*. The difficulty in comparing small differences in large quantities.
- *Choice-supportive bias*. The tendency to remember one’s choices as better than they actually were.

## High-dimensionality Blindness

The limitations of human intelligence related to high-dimensional visualization, perception, attention, etc. are well known. The total impact is that human intelligence is “blind” to analyze phenomena that have many interactive factors and beyond some level of complexity. Some of the key contributors to this type of weakness are given below (Kordon, 2020):

- *3D visualization limit*. The constraints of human intelligence on imagining and identifying patterns beyond three dimensions.
- *Multivariate relationships limit*. The difficulties of human intelligence in discovering, defining, and understanding the dependencies among many factors (variables).
- *Attention limit*. The limitations of human intelligence in paying attention to or tracking a large number of factors.
- *Combinatorial limit*. Human intelligence is not very impressive in tasks that require evaluation of many alternative solutions. A clear example is the success of AI in beating humans in complex games, such as chess and recently the board game Go.

## **Uncontrolled Emotional Intelligence**

Probably the best-known limitation of human intelligence is that rationality of human thought is entrapped in the emotional “gravity” of human behavior. If a person cannot handle the negative influence of her/his emotions, the consequences can be damaging, especially making important business decisions. Fortunately, the broad area of emotional intelligence delivers a popular framework and methods for controlling human emotions and reducing the effect of emotionally-driven responses and decisions. The classical book on this topic is (Goleman, 1995).

## **Low Decision-making Speed**

Human intelligence depends on the way our brain operates. Unfortunately, the brain’s processing speed is very slow in comparison with contemporary computers. In addition, its memory capabilities are not impressive and degrade with age or illness. Both of these limitations become an important obstacle when optimal business decisions, based on millions of records, hundreds of diverse factors, and thousands of conditions have to be made in seconds in real time.

## **Performance Swings**

Another well-known weakness of human intelligence is its changing and unpredictable performance. Even the best and the brightest have their bad times due to a number of reasons. In general, performance swings are a result of the biological limitations of human bodies.

## **KEY BENEFITS OF HUMAN INTELLIGENCE FROM AI**

The majority of people in businesses look at the capabilities of AI as a unique opportunity to compensate for these limitations and to enhance human intelligence and improve productivity. The term “augmented intelligence” has even been used to define this cooperation between human intelligence and AI technologies. A search engine can be viewed as an example of augmented intelligence (it augments human memory and factual knowledge), as can natural language translation (it augments the ability of a human to communicate).

Most of the weaknesses of human intelligence can be compensated for by augmented intelligence. Cognitive biases can be reduced or eliminated by using model-based decisions. AI gives human intelligence the missing high-dimensionality capabilities, as well as the ability to make fast, emotionless decisions using big data.

Most of these systems raise significantly the performance level of all users and make it consistent all the time.

The following benefits of human intelligence from AI will be discussed shortly below:

## **Decisions Based on High Dimensionality**

AI augments human decision-making with conclusions based on dependences of millions of factors and high complexity. The unprecedented ability of AI-based algorithms to manipulate the large amounts of data each of us generates online enables people to organize and use this information in staggering new ways, affecting all spheres of our existence. Gradually, business decisions are based on AI complex algorithms that take into account various factors that relate to a specific business topic. A typical example is the financial industry where seventy percent of all financial transactions are already done by algorithms (Coleman, 2019).

## **High Accuracy on Specific Tasks**

AI delivers higher precision than humans in problems related to visual analysis of interactive factors. Typical cases are some types of medical diagnostics. For example, AI is already helping to digitize information and speed up analysis of data in pathology, and, when teamed with a pathologist, is showing an 85 percent reduction-in-error rate in diagnosing cancers. In the United States, AI is improving the accuracy of breast cancer screenings. AI is already better than humans at predicting in vitro success (Coleman, 2019).

Another example of high accuracy solution is using factors like timing of holidays and medical leaves taken, and even how quickly employees answer emails—factors identified with machine learning models—AI can now predict with 85 percent accuracy whether someone will leave within three months (Davenport, 2018).

## **AI Increases Existing Average Knowledge at Expert's Level**

Most of AI solutions are based on the extraction of the best available knowledge from experts and available data. This new AI-generated knowledge is communicated across the business and is accessed by all interested stakeholders. As a result, the difference between non-experts and experts in understanding the processes, factors, dependencies, and patterns in the business is significantly reduced. This lowers the risk of making inappropriate decisions and protects the business from a knowledge gap if experts leave the company.

## **Repetitive Cognitive Tasks Can Be Carried Out With the Help of AI**

With the increasing capabilities of AI, more repetitive mental activities have been transferred from human to artificial intelligence. Typical cases are scanning and pre-processing job applications in human resources, low-level customer service in many big retailers, and preparing some routine tasks in accounting. AI methods will be more widely used in more accurate and complex decision-making by all levels in business hierarchy.

## **KEY WEAKNESSES OF AI**

Some of the observed disadvantages of AI demonstrated in real-world applications will be discussed, such as:

### **Limitations of “Narrow” AI**

AI is divided in two types – “narrow” or “weak” AI and “general” or “strong” AI. “Narrow” AI can intelligently perform specific tasks better than a human, but lacks what we think of as common sense and the kind of comprehensive intelligence that belongs to human beings. The level of narrow-based AI varies significantly and some of the “not-so-intelligent” applications are challenged by humans. For example, the most successful applications of “narrow” AI - speech recognition and object recognition aren’t intelligence, they are just pieces of intelligence. For real human-like intelligence we also need reasoning, language, and analogy, none of which is nearly well handled by the “narrow” AI.

“Strong” AI, or Artificial General Intelligence (AGI) refers to a machine that will have an authentic capacity to “think” and will be capable of performing most human tasks. General intelligence AI will require both mechanisms like deep learning for recognizing images and machinery for handling reasoning and generalization, closer to the mechanisms of classical AI and the world of rules and abstraction. Unfortunately, the most-powerful current technology of “narrow” AI - deep learning has difficulties to incorporate background knowledge, which is critical for AGI functionality. The point is that simply creating larger and deeper neural networks and larger data sets and bigger machines is not enough to create the needed human-level AI.

The discrepancy between the limited, even impressive, capabilities of “narrow” AI in performing specific human tasks and the expectation for human-like performance, which could be delivered in the future by the AGI, but is unavailable now, creates disappointment to many potential users of the technology.

## **Insufficient Trust**

At the basis of the trust issue with current AI applications is the black-box nature of most of the developed solutions, especially when using powerful algorithms, such as deep learning. These models are not transparent and it is difficult to explain the generated results and machine recommendations. As a result, some users are not convinced in the AI-generated decisions. The more an AI system leverages machine learning and neural networks to crunch immense amounts of data, the less likely it is for a human to understand how the AI has arrived at its conclusion. MIT's Technology Review magazine called this model opacity "the dark secret at the heart of AI" (Mitchell, 2019). The fear is that if we don't understand how AI systems work, we can't really trust them or predict the circumstances under which they will make errors.

The ultimate level of lack of trust is the growing concern that black-box models can be "cheated" and used inappropriately. As one example, a group from the University of California at Berkeley designed a method by which an adversary could take any relatively short sound wave—speech, music, random noise, or any other sound—and perturb it in such a way that it sounds unchanged to humans but that a targeted deep neural network will transcribe as a very different phrase that was chosen by the adversary (Mitchell, 2019).

## **Algorithmic Biases**

Contrary to the expectation of pure objectivity of model-based decisions, it has been observed in many business applications that AI has its own biases. They are based on the tendency of machine learning algorithms to produce inappropriately biased models about loans, housing, jobs, insurance, parole, sentencing, college admission, and so on. The likely root causes of algorithmic bias lie in the data rather than in the deliberate malfeasance of corporations (Russel, 2019). Especially sensitive to the unequal distribution of the factors in the training data are the models based on deep learning. Usually poorly collected data with clear dominance of some factors lead to models with biased results magnifying the influence of certain factors at the expense of others. A recommended solution to this widespread problem is to counteract algorithmic biases with data sources diversity, especially in marketing (Stern, 2017). Another remedy is to balance individual models' biases by combining different algorithms in an ensemble. Various model ensemble-building techniques are described in (Kordon, 2020).

## **Lack of Creativity**

The issue of “narrow” AI creativity needs some explanations. On the one hand, many AI algorithms, especially evolutionary computation, create novel results. Of special importance to business applications is one specific algorithm – genetic programming, which has been used as patent generator of electronic circuits. There are many other examples of creating innovative solutions in design, new patterns discovery, and art (Kordon, 2020). On the other hand, “narrow” AI cannot recognize by itself the creative nature of the novel solutions it has generated. As a result, “narrow” AI creativity needs to be recognized as such by human intelligence. The hope is that AGI will be autonomously creative and being able to understand and judge what it has created without human intervention. Until then the burden of proof of novelty is on human intelligence.

## **KEY NEGATIVE EFFECTS OF AI ON HUMAN INTELLIGENCE**

The following destructive impact of AI will be considered:

### **Mental Laziness**

One of the most dangerous consequences of the growing influence of AI technologies is reducing cognitive activities by blindly relying on them in almost any action. In some cases, such as following driving instructions derived by complex AI-based systems, is a big advantage and make perfect sense. However, even in this case, drivers should not keep on directions that are not appropriate according to her/his knowledge and experience. Fast access to various types of knowledge by search engines, based on AI algorithms, is of tremendous benefit to any cognitive activity. The problem is when using the power of AI is at the expense of improving human intelligence. The growing desire to “outsource” thinking to AI can lead to dangerous and long-term effects on the way humans think and analyze. The mindset that AI will make better decisions based on big data and relieve human intelligence from this burden may lead to gradual mental degradation and dependence on machine intelligence. However, in this addiction, we should be careful of the potential of machine intelligence to behave inadequately, even stupid. Machine stupidity creates a tail risk related to low probability occurrences. Machines can make many good decisions in 99% of the time and then one day fail spectacularly on a tail event that did not appear in their training data. This can be expected from “narrow” AI due to the limited nature of data collection it uses for its development (Mitchell, 2019).



## **Job Insecurity**

Probably the most negative effect of AI is the growing concern of massive restructuring of the labor force towards high-tech jobs. Jobs with low educational level and routine operations are the first that could be replaced by AI systems. This may contribute to stress and disappointment of many current employees in this category. It is strongly recommended that they take actions for training in the new skillset, required for AI systems as soon as possible.

## **Enhanced Destructive Capabilities at a Society Level**

This includes activities, such as: hacking, fake news, and privacy invasion. All of these negative consequences of the AI invasion directly influence human intelligence at the individual level. Of special concern are cases of hacking medical information, derived by AI. For example, it has been demonstrated a possible adversarial attack on deep neural networks for medical image analysis: they showed that it is not hard to alter an X-ray or microscopy image in a way that is imperceptible to humans but that causes a network to change its classification from, say, 99 percent confidence that the image shows no cancer to 99 percent confidence that cancer is present (Mitchell, 2019).

Recently, the capabilities for extracting private information and for future manipulation become very sophisticated due to AI. Some chatbots are reportedly designed to feign romantic interest in order to trick unsuspecting internet users into revealing personal information (such as childhood pets and parent names) that might be used to gain access to bank accounts, credit cards, and the like (Smith, 2018).

The biggest issue at a society level is that Big Business as well as Big Government is collecting detailed information about everything we do so that they can predict our actions and manipulate our behavior. Big Business and Big Government monitor our credit cards, checking accounts, computers, and telephones, watch us on surveillance cameras, and purchase data from firms dedicated to finding out everything they can about each and every one of us (Smith, 2018).

## **RESISTANCE OF HUMAN INTELLIGENCE TOWARDS AI**

For some people, the new, powerful capabilities of AI technologies are a direct threat to their jobs and current competencies. They look at AI as a negative amplifier, i.e., it enhances their weaknesses and ignorance, and raises the chance of decreased performance due to new skillset gaps. This attitude may evolve into an AI-driven inferiority complex with different forms of resistance, such as questioning the

recommended decisions and results, exaggerating cases with wrong predictions, incorrectly using deployed solutions, or not using them at all.

## **Skepticism Towards AI Capabilities**

Skepticism is usually the initial response of the final users of the AI-based technology on the business side. Several factors contribute to this behavior, such as lack of awareness of the technical capabilities and application potential of AI, lessons from other overhyped technology fiascos in the past, and caution about ambitious high-tech initiatives pushed by management.

Skepticism is a normal attitude if risk is not rewarded. Introducing emerging technologies, such as AI requires a risk-taking culture from all participants in this difficult process. The recommended strategy for success and reducing skepticism is to offer incentives to the developers and the users of the technology.

## **Factors for Resistance Towards AI**

Analyzing the key factors that contribute to the resistance of human intelligence towards AI in many real-world applications, are discussed briefly below (Kordon, 2020).

- *The business fails to reinforce AI-based applications properly.* The leadership has not sent a clear message to all employees that AI will not endanger their jobs, and encouraged users of new technology with corresponding benefits, proportional to the productivity gain.
- *Business users are not prepared to change their behavior.* Adopting new processes and tools can disturb even the most enthusiastic employees. It is strongly recommended that companies deploying AI solutions should provide constant training and coaching of employees not only in how to use the new technology but also in understanding the implications for decision-making at every level of the company. Without this, the common response is to reject and resist change with all possible means available.
- *Experts are not committed to using AI-based solutions.* Business experts weren't engaged from the beginning of the company's raids into AI-driven solutions or they don't see the value of the technology. In either case, it is likely that the business leadership has not communicated the company's vision of how the experts and their careers would benefit from AI.
- *Data scientists and business teams are not communicating effectively.* Too often, data scientists "throw the models and run" and let the users in the business struggle with all the pains of the deployment of the models. That

approach rarely works, and triggers complaints from the users. Most of them will resist any future attempts to implement such technology.

- *The delivered solutions are not user-friendly.* A specific issue in this category is the case of “black-box” or purely academic AI-based solutions. They provide clunky, overly complicated insights that are impossible for business users to fully understand. As a result, users do not trust this “academic abracadabra” and question its application.

Most of the above factors contributing to the resistance toward AI-based solutions can be resolved with an improved business strategy, proper incentives, communication, and project management.

## **HOW TO IMPROVE THE INTERACTION BETWEEN HUMAN AND ARTIFICIAL INTELLIGENCE**

Some suggestions for enhancement of the collaboration between human and artificial intelligence will be discussed with examples from industrial applications:

### **Model-based Decision-making**

An important generic area of improving human and artificial intelligence interactions is by gradually using AI-generated models in business decisions. The value of changing the business culture by replacing qualitative “guts-based” business rules with quantitative “AI-model-based” decisions will be illustrated.

One of the key reasons for using model-based decisions is that only recently the businesses have an opportunity to obtain broad access to high-volume data from a variety of sources and analyze it with very sophisticated machine learning algorithms. The hypothesis is that these two factors, when combined, dramatically increase the probability of defining something like “truth” and do so in a short time window that is actionable. This allows businesses to make more correct decisions (based on “truth”) and have a competitive advantage over those competitors using heuristic-based decisions (based on gut feeling).

Some data supports this hypothesis. According to a recent survey, companies in the top third of their industry in the use of data-driven model-based decision-making are, on average, 5% more productive and 6% more profitable than their competitors (McAfee and E. Brynjolfsson, 2012).

Advantages of AI model-based decisions:

- *Objective nature of decisions.* Mathematical models represent objective relationships among the factors influencing a business problem. Using their predictions, with their confidence limits, in the decision-making process increases the unbiased nature of the process. Most business rules are still defined subjectively by experts. The quantitative metrics in them, however, are based on objective relationships derived by machine learning algorithms.
- *Decisions based on multivariate factors.* Many AI-model-based decisions use predictions from multivariate approaches or models that include many factors influencing the business problem. These decisions are more adequate to the complexity of real-world problems.
- *More accurate decisions.* Model-based decisions transform estimates and forecasts into much more precise, accurate predictions that remove enormous chunks of risk from the decision-making process. This substantial increase in accuracy translates into higher degrees of confidence throughout an organization, which in turn drives fundamental changes in behavior and high trust in the overall decision-making process.
- *Quantitative estimates about the future.* A big advantage of predictive and especially forecasting models is their numerical most probable forecast, with its low and high confidence limits over a specific forecasting horizon. It allows one to specify the “shape of the future” in quantitative terms and narrows down the uncertainty in making the business decision.
- *Improved existing business rules.* The most popular approach in using developed models in decision-making is to enhance existing business rules with more accurate predictions. This is the best way to integrate existing problem knowledge, based on previous experience, with more objective quantitative information, based on discovered complex relationships between the factors influencing the problem.

An interesting example of the benefits of model-based decision is the use of AI-based algorithms as a board member. In May 2014 Deep Knowledge Ventures – a Hong Kong venture-capital firm specializing in regenerative medicine – broke new ground by appointing an algorithm called VITAL to its board. VITAL makes investment recommendations by analyzing huge amounts of data on the financial situation, clinical trials and intellectual property of prospective companies. Like the other five board members, the algorithm gets to vote on whether the firm makes an investment in a specific company or not (Harari, 2017).

## **Introducing a Model-based Decision-making Process**

Another key step in building a model-based decision culture in a business is defining a work process for consistent implementation of this approach across the organization. The best-case scenario is where this type of decision-making is integrated into existing work processes. A good candidate is the popular Sales and Operations Planning (S&OP) process (Wallace, 2011). The S&OP process includes an updated forecast that leads to a sales plan, a production plan, an inventory plan, a customer lead time (backlog) plan, a new product development plan, a strategic initiative plan and a resulting financial plan. The frequency of the planning process and the planning horizon depend on the specifics of the industry. Although invented over 20 years ago, S&OP is currently in a stage of wide popularity and rapid adoption, especially with AI-based technology growth.

In order for the integration of model-based decision-making into these work processes to be successful, the following principles are strongly recommended (Kordon, 2020):

- *Balancing stakeholders' interests.* Understanding the stakeholders' interests and concerns is a winning strategy for avoiding potential political issues. For example, typical business clients for AI-based modeling projects expect trustworthy results that are similar to their own expert judgment and can help them make correct decisions and perform profitable planning. They prefer simple explanations and consistent performance. Unfortunately, it is a challenge with the machine learning generated black-box models. Their main concern is trying to avoid a decision fiasco and its negative impact on their careers.
- *Handling biases.* One of the realities of applying model-based decisions is that users have preliminary opinions about how the future will look based on their knowledge. Handling this biased vision is one of the biggest challenges in managing model-based predictive projects. The most widespread bias is based on overconfidence in the power of modern AI-based algorithms and ignoring their limitations. Most people are overly optimistic in their forecasts while they significantly underestimate future uncertainty. Often, people replace forecasting with their own wishful thinking.

Some subject matter experts are biased toward their hypotheses about potential process or economic drivers. If their list is not supported by the data and their favorite drivers have not been selected by the AI algorithms used, they want detailed explanations. It is strongly recommended that the statistical basis of the variable selection should be described carefully and the supporting data be shown. In some

situations, it is possible to reach a compromise and include the expert-defined drivers in the inputs used for modeling and let the AI-based selection algorithms do the final selection.

- *Avoiding political overrides.* In some cases, AI-based forecasts are corrected due to clear departmental political purposes. One extreme is sandbagging, when some sales departments lower the statistical forecast to reduce their sales quota in order to guarantee their bonuses. The other extreme is sales departments that have losses due to back orders. They prefer to raise the AI-based forecast in the hope of managing inventory levels via the sales department is forecast.

An example of a successful implementation of AI-model-based decision-driven cultural change is the large -scale raw materials forecasting project at Dow Chemical (Kordon, 2012). As the largest US chemical company with a broad range of products, Dow Chemical is using many different raw materials. Their total cost is in the range of billions of dollars. Reducing this cost with reliable forecasting models can deliver significant value and directly influence the company's profit. In order to accomplish this goal, a large-scale forecasting project for the development and deployment of predictive models for the top 51 costliest raw materials prices was initiated with the support of the purchasing department.

The project included several different teams, related to selected raw materials prices in the corresponding geographic regions. Each team involved the most experienced experts familiar with the specific requirements of price negotiation and with the purchasing history. The key objective of the first step of the project, entirely driven by human intelligence, is to define the list of potential economic drivers related to the raw materials prices. In several brainstorming sessions, the experts recommended three types of driver. The first type included macroeconomic factors, such as Gross Domestic Product (GDP), consumer confidence, population growth, and exchange rates. The second type included microeconomic data from the supply and demand sectors of the economy related to the raw materials prices, for example the construction index and personal care index. The third type included competitive or related prices of other raw materials or products.

Then AI took the lead and did automatic selection of the most important economic drives and develop 51 forecasting models for all 51 raw materials of interest. The most important part of the project, however, was the effort to integrate the forecasting models into the existing business decision rules. In order to accomplish this goal, a work process for using the forecasting models in price negotiation was proposed, discussed, and defined. It included automatic data collection, model execution, report generation, and instructions on how to use the AI-based forecasts.

## **Comparing Expert's Judgement With AI-based Models' Predictions**

A key challenge in improving human and artificial intelligence interaction is to build trust in AI-generated solutions. A recommended approach is to organize an objective and fair performance assessment to help human intelligence to realistically evaluate the advantages and disadvantages of AI in specific activities. Of special importance is a comparison with existing heuristic-based decisions. An example of such a comparison with the discussed raw materials forecasting project is given below:

The objective of the study was, for the same forecasting horizon of six months, to compare the forecasting errors between the developed AI-generated predictive models and the judgmental forecast given by the experts based on the best of their knowledge. The performance metrics is the mean absolute percentage error (MAPE) between the forecasted and real raw materials prices during this six months trial period. A summary of the results show, that the range of the 22 deployed and used models MAPEs is between 1.7% (the best model) and 28% (the worst model) with a mean error of 10.3% (which is on the borderline of a "good" forecast). The corresponding summary results for the experts' judgmental forecasts, demonstrate an inferior performance, with a broader error range between 2.7% (the best model) and 39.3% (the worst model) and a much higher mean MAPE of 16.3%. This clear advantage of statistical versus heuristic-based forecasting was a critical factor in convincing even the biggest skeptics about using the benefits of model-based decision making for more profitable raw materials price negotiation in purchasing (Kordon, 2012).

## **Combining Human and Model-based Decisions**

The most likely scenarios for combining human and artificial intelligence are those which have solutions that are in some 'shades of gray'. One scenario is when the accuracy of the AI models needs human confirmation. Another option is where there is the need to add more context in the models, particularly around human interaction or human understanding. A typical scenario is where AI-generated model is used for accurate prediction but the allocation or decision regarding resources for executing the solution may be subjective at some level (e.g. selecting the proper raw materials price during price negotiation in the case of the discussed project.) The more complicated the scenario and the greater the degree of 'shades of gray', the greater are the odds for a well-balanced partnership between human and artificial intelligence.

Some business applications will involve 80 percent AI-based decisions and 20 percent human ones; others the opposite. Systematic design activity is necessary to determine how humans and machines will augment each other's strengths and

compensate for their weaknesses. At the investment firm Vanguard, for example, a new Personal Advisor Services (PAS) offering combines automated investment advice and guidance from human advisors at a lower cost than purely human-advised investing. The PAS technology performs many of the traditional tasks of investment advising, including constructing a customized portfolio, rebalancing portfolios over time, tax loss harvesting, tax-efficient investment selection, and creating recommendations for safe withdrawal amounts for retirees. The system took over some tasks from advisors, including acquiring basic information from customers and presenting financial status information to them—tasks that were sometimes considered tedious for human advisors anyway (Davenport, 2018).

## **CONCLUSION**

The importance of analyzing the interaction between human and artificial intelligence is growing with the fast pace of applying AI-based systems in industry. Some ideas for such an analysis that can help in improving the efficiency of AI in business applications are discussed in the chapter. A step in this process is defining of the limitations of both human and artificial intelligence that could be balanced by better integration. The focus is on the following limitations of human intelligence: cognitive biases, high-dimensionality blindness, uncontrolled emotional intelligence, low decision-making speed, and performance swings. On the side of artificial intelligence, the focus is on the following boundaries of the technology: the nature of the “narrow” AI, insufficient trust due to the black-box nature of machine learning models, algorithmic biases based on limited training data used in model development, and the lack of creativity. The other recommended step is in evaluating the positive and negative effects of AI on human intelligence in business applications. The following generic benefits of AI on human intelligence are considered: making decisions based on many inter-dependent factors, increasing the accuracy of some specific tasks, especially in medical diagnostics, potential for raising the average specific skillset to expert’s level, and outsourcing some specific routine cognitive tasks to AI. The key negative effects of AI on human intelligence are defined as: potential for mental laziness due to reducing cognitive activities by blindly relying on AI, growing job insecurity, and some negative activities at a society level such as: hacking, fake news, and privacy invasion. The ultimate form of negative effect of AI over human intelligence is growing skepticism leading to resisting the technology. Some factors that contribute to this effect and should be taken into account are identified.

Improving the interaction between human and artificial intelligence by using model-based decision-process is recommended in the chapter. It takes into account the factors in the analysis and is of big practical importance to many business



applications. The advantages of the proposed approach are illustrated with a large-scale industrial application with many users. The model-based decisions by using AI-generated forecasts have been significantly more accurate than the “guts”-based decisions based on human intelligence only.

## REFERENCES

- Clarke, R. (2019). Principles and processes for responsible AI. *Computer Law & Security Review*, 35(4), 410–422. doi:10.1016/j.clsr.2019.04.007
- Coleman, F. (2019). *A human algorithm: how artificial intelligence is redefining who we are*. Counterpoint.
- Davenport, T. H. (2018). *The AI advantage: how to put the artificial intelligence revolution to work*. MIT Press. doi:10.7551/mitpress/11781.001.0001
- Goleman, D. (1995). *Emotional intelligence*. Bantam Books.
- Harari, Y. N. (2017). *Homo deus: a brief history of tomorrow*. Harper. doi:10.17104/9783406704024
- Human intelligence definition. (n.d.). Retrieved from [https://en.wikipedia.org/wiki/Human\\_intelligence](https://en.wikipedia.org/wiki/Human_intelligence)
- Kordon, A. (2012, October). *Applying data mining in raw materials forecasting*. Paper presented at the SAS Analytics 2012 Conference, Las Vegas, NV.
- Kordon, A. K. (2020). *Applying data science: how to create value with artificial intelligence*. Springer.
- List of cognitive biases. (n.d.). Retrieved from [https://en.wikipedia.org/wiki/List\\_of\\_cognitive\\_biases](https://en.wikipedia.org/wiki/List_of_cognitive_biases)
- Marcus, G., & Davis, E. (2019). *Rebooting AI: building artificial intelligence we can trust*. Pantheon Books.
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, (October), 2012. PMID:23074865
- Mitchell, M. (2019). *Artificial intelligence: a guide to thinking human*. Farrar, Strauss and Giroux.

***Practical Issues in Human and Artificial Intelligence Interaction***

Russel, S. (2019). *Human compatible: artificial intelligence and the problem of control*. Viking.

Smith, G. (2018). *The AI delusion*. Oxford University Press. doi:10.1093/oso/9780198824305.001.0001


Stern, J. (2017). *Artificial intelligence for marketing*. John Wiley & Sons. doi:10.1002/9781119406341

Wallace, T. (2011). *Sales and operations planning: beyond the basics*. T.F. Wallace & Company.

## Chapter 4

# Policy and Management Issues of Artificial Intelligence

**Bistra Konstantinova Vassileva**

 <https://orcid.org/0000-0002-5976-6807>  
*University of Economics, Varna, Bulgaria*

### ABSTRACT

*The capacity for AI research, technology, and application is seen as vital to national competitiveness, security, and economic strength. In the last few years, several countries and regions have developed and released AI strategic plans, thus setting up a race to become the global leader in the field. The chapter starts with an overview of the latest development in AI legislation and governance principles. The first section begins with a review of available policies and strategies on AI by countries and regions. Some best practices in AI governance are presented as well. The specifics of AI ecosystems are discussed in the second section. Gephi software tool is used to visualize the mapping of the Italian AI ecosystem. The chapter ends with conclusions and recommendations aimed at the future development of policy and management for responsible AI implementation.*

### INTRODUCTION

The capacity for AI research, technology, and application are seen as vital to national competitiveness, security, and economic strength. In the last few years, several countries and regions have developed and released AI strategic plans, thus setting up a race to become the global leader in the field (Olley, 2020:14). The chapter starts with an overview of the latest development in AI legislation and governance principles. The first section begins with a review of available policies and strategies

DOI: 10.4018/978-1-7998-4285-9.ch004

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

on AI by countries and regions. Some best practices in AI governance are presented as well. The specifics of AI ecosystems are discussed in the second section. Gephi software tool<sup>1</sup> is used to visualize the mapping of the Italian AI ecosystem. The chapter ends with conclusions and recommendations aimed at the future development of policy and management for responsible AI implementation.

## **BACKGROUND**

Globalization has changed the strategic context for business and nowadays it is viewed not only as a geographical expansion, but rather as a new operating theory of the world based on connectedness among pre-existing political, social, economic, thematic, and geographic boundaries (Singer, 2006: 51). Both the connectedness and complexity have become a source of instability and risk, as well as a driver for accelerating the reorganization of the global economic landscape. More or less, complexity has become the new norm for business, requiring a new perspective<sup>2</sup>. Three main schools of thought can be distinguished in the field of complexity. These schools of thought originate from three different academic institutions: the Free University of Brussels, the University of Stuttgart, and the Santa Fe Institute. The concept of complexity was initially developed in Brussels and Stuttgart by chemists and physicists working on scientific topics connected with emergent structures and disequilibrium dynamics. The concept was further elaborated by researchers from the Santa Fe Institute in the late 1980s with a broad focus on and implication for the economy. In the sense of complexity, economy was defined not as something given and existing, but something which is formed from a constantly developing set of institutions, arrangements, and technological innovations (Arthur, 2013). According to the “Santa Fe” perspective<sup>3</sup>, complexity consists of six characteristics which are presented in Table 1.

The essence of complexity, according to Arthur (2013), is about the formation of structures and how this formation affects the objects causing it. Over the years it has been examined in different economic and cultural contexts, e.g. the “Austrian perspective” (Montgomery, 1999). Based on the abovementioned notions and views about complexity, a conclusion could be drawn that a system is complex when (1) it is an open system, (2) its behavior crucially depends on the details of its parts (Parisi, 1999:560), (3) it is a higher-order structure, arising from a set of lower level structures, (4) it may exhibit behaviors that are emergent, and (5) its relationships are nonlinear and contain feedback loops.

Following the characteristics of the complex adaptive systems Beinhocker (cited in Gintis, 2006:1) suggested that the market economy follows an evolutionary dynamics, hence the analytical tools applied in evolutionary biology can be used to

*Table 1. Main characteristics of complexity*

Characteristic	Description
Interaction	Dispersed interaction among heterogeneous* agents acting locally on each other in some space.
Controller	No global controller that can exploit all opportunities or interactions in the economy even though there might be some weak global interactions.
Organization	Cross-cutting hierarchical organization with many tangled interactions.
Adaptation	Continual adaptation by learning and evolving agents.
Innovation	Perpetual novelty as new markets, technologies, behaviors, and institutions create new niches in the “ecology” of the system.
Dynamics	Out-of-equilibrium dynamics with either zero or many equilibria existing and the system unlikely to be near a global optimum.

\* some Santa Fe Institute models assume dispersed homogeneous agents

Source: Adapted by Arthur, W. B., Durlauf, S. N., and Lane, D. A. (1997). Introduction, in: *The Economy as an Evolving Complex System II*. Arthur, W. Brian, Steven N. Durlauf and David A. Lane, eds. Reading: Addison-Wesley, pp. 3–4.

analyze institutions, firms, technology and business culture. This conceptualization is supported by Arthur (2013:1) who holds the opinion that “complexity economics builds from the proposition that the economy is not necessarily in equilibrium: economic agents (firms, consumers, investors) constantly change their actions and strategies in response to the outcome they mutually create.” The intense penetration of AI in industry and human lives raises the importance of the ability of governments, institutions and companies to adapt their strategies and management practices to complexity.

## MAIN FOCUS OF THE CHAPTER

### AI Policies and Strategies: An Overview

During the last few years various institutions and organizations at different hierarchical levels focus their attention on developing policies, strategies, plans, programs, etc. to establish and to stimulate AI capabilities at national level to power their future, digital economy. These initiatives are usually driven by government-wide partnerships comprising diverse types of entities.

The available policies and strategies on AI are reviewed by countries and regions. They generally call for more investment to build the AI workforce and research and development capacity. The main focus of studied strategic documents is placed on the AI effects on jobs and economies as well as on the social, economic, and ethical

## ***Policy and Management Issues of Artificial Intelligence***

implications of AI. AI policies developed as part of these national strategies vary widely from country to country, but there are several common elements: governance and regulation, ethics, security, research, and people.

The AI strategies of different states focus on different aspects of AI policy such as scientific research, talent development, skills and education, public and

*Table 2. An overview of AI policies and strategies by countries, 2017*

<b>Time Period</b>	<b>Country</b>	<b>Name of the Document</b>	<b>Short Description</b>
March, 2017	Canada	Pan-Canadian AI Strategy	The strategy has four goals: (1) to increase the number of AI researchers and graduates, (2) to establish three clusters of scientific excellence, (3) to develop thought leadership on the economic, ethical, policy, and legal implications of AI, and (4) to support the national research community on AI.
March, 2017	Japan	AI Technology Strategy	It is based on the Industrialization Roadmap which final goal by 2030 is the creation of ecosystems built by connecting multiplying domains.
May, 2017	Singapore	AI Singapore announced	The National Research Foundation (NRF) Singapore will launch a national programme in AI to catalyse, synergise and boost Singapore's AI capabilities to power future, digital economy (AI.SG). Three objectives: 1/ Use AI to address major challenges that affect society and industry; 2/ Invest in deep capabilities to catch the next wave of scientific innovation; 3/ Broaden adoption and use of AI and machine learning within industry.
July, 2017	China	New Generation AI Plan	Comprehensive plan with initiatives and goals for R&D, industrialization, talent development, education and skills acquisition, standard setting and regulations, ethical norms, and security
October, 2017	UAE	AI Strategy 2031	The strategy is intended to achieve the UAE Centennial 2071 objectives, boost government performance, and invest in AI adoption. The UAE's AI strategy covers development and application in nine sectors: transport, health, space, renewable energy, water, technology, education, environment, and traffic.
October, 2017	Finland	AI Strategy (Finland's Age of Artificial Intelligence)	It will strive to: (1) increase the competitiveness of business and industry; (2) provide high-quality public services and improve the efficiency of the public sector; (3) ensure a well-functioning society and wellbeing for its citizens.
December, 2017	China	Three-Year Action Plan	It advances four major tasks: (1) focus on developing intelligent and networked products such as vehicles, service robots, and identification systems, (2) emphasize the development AI's support system, including intelligent sensors and neural network chips, (3) encourage the development of intelligent manufacturing, and (4) improve the environment for the development of AI by investing in industry training resources, standard testing, and cybersecurity

Source: Author's work

private sector adoption, ethics and inclusion, standards and regulations, and data and digital infrastructure. The main building blocks of these strategic documents usually include ideas or innovation, people, infrastructure, business environment, places/communities. Governments' vision is aimed at developing relevant capacities in their countries to reach the leading position of the artificial intelligence and data revolution. Some of these documents include a roadmap (e.g. Japan) or action plan to achieve the planned strategic goals.

At the EU level the implementation of national policies and regional programmes is vital to complement European coordinated actions (Delponte, 2018:25) because of the general feeling that Europe is falling behind in the race for AI leadership strong despite the solid AI basic research which is done there. The Finnish government has established an AI ethics committee to gain understanding on ethical principles and to ensure that Finland's AI development is human-oriented and based on trust. Policies directed to the development of ethical guidelines engage both companies and public administration. The government has also announced a new commission to investigate how AI and algorithmic decision-making will affect society. The UK Artificial Intelligence Sector Deal is the first commitment from government and industry to realize this technology's potential, outlining a package of up to £0.95bn of support for the sector, which includes government, industry and academic contributions. The vision of how the UK can respond to the broader opportunities and challenges for society posed by AI includes the ambition of leading the world in the safe and ethical use of data through a new Centre for Data Ethics and Innovation. The National Institution for Transforming India (NITI Aayog) identifies healthcare, agriculture, education, smart cities, and smart mobility as the priority sectors that will benefit the most socially from applying AI. This organization recommends setting up a consortium of Ethics Councils at each centre of research excellence and the international centre of transformational AI (ICTAI), developing sector specific guidelines on privacy, security, and ethics, creating a National AI Marketplace to increase market discovery and reduce time and cost of collecting data, and a number of initiatives to help the overall workforce acquire skills. Strategically, the government wants to establish India as an "AI Garage," meaning that if a company can deploy an AI in India, it will then be applicable to the rest of the developing world.

## **Specifics of AI Ecosystems**

As it was described in the first section of this chapter, the twenty-first century revolution is characterized by complex nonlinear emergent behaviors (Hamel, 2002). Digitization is causing dramatic changes in companies' business ecosystems, making them larger and more complex to manage strategically, due to the greater number and more diverse CSFs and KPIs. As ecosystems enable companies to respond

## Policy and Management Issues of Artificial Intelligence

*Table 3. An overview of AI policies and strategies by countries, 2018-2019*

Time Period	Country	Name of the Document	Short Description
January, 2018	Kenya	Blockchain and AI Task Force	Blockchain & Artificial Intelligence taskforce will provide the roadmap to contextualize on the application of these emerging technologies in the areas of financial inclusion, cybersecurity, land titling, election process, single digital identity and overall public service delivery.
January, 2018	Taiwan	AI policies and developments in Taiwan	The first overall policy that is relevant for AI is the 5+2 Innovative Industries Plan. The second overall policy that is relevant for AI is the Forward-Looking Infrastructure Development Program. Aside the general policies described above, the Taiwanese government also published specific AI policies.
January, 2018	Denmark	Strategy for Digital Growth	The strategy has three goals: (1) make Danish businesses the best at using digital technologies; (2) have the best conditions in place for the digital transformation of business; and (3) ensure every Dane is equipped with the necessary digital skills to compete.
March, 2018	Italy	White Paper on AI (AI at the Service of Citizens)	It exclusively focuses on how the government can facilitate the adoption of AI technologies in the public administration.
March, 2018	France	AI Strategy (AI for Humanity)	The strategy focuses on (1) an economic policy based on data; (2) towards agile and enabling research; (3) anticipating and controlling the impacts on jobs and employment; (4) using AI to help create a more ecological economy; (5) the ethics of AI; (6) for inclusive and diverse AI.
April, 2018		First Workshop for Strategy	The human rights framework was described as an aspirational roadmap and moral compass for actors in the AI space.
April, 2018	UK	AI Sector Deal	The Industrial Strategy is built on 5 foundations: 1/ Ideas - the world's most innovative economy; 2/ People - good jobs and greater earning power for all; 3/ Infrastructure - a major upgrade to the UK's infrastructure; 4/ Business environment - the best place to start and grow a business; 5/ Places - prosperous communities across the UK.
April, 2018	EU	Communication on AI	EU Commission aims to (1) increase the EU's technological and industrial capacity and AI uptake by the public and private sectors; (2) prepare Europeans for the socioeconomic changes brought about by AI; and (3) ensure that an appropriate ethical and legal framework is in place.
May, 2018	Australia	Australian Budget	Development of AI in Australia will be supported by the budget. The government will create a Technology Roadmap, a Standards Framework, and a national AI Ethics Framework to support the responsible development of AI.
May, 2018	USA	White House Summit on AI	The United States' national strategy on AI is a concerted effort to promote and protect national AI technology and innovation. It directs the Federal government to pursue five pillars for advancing AI: (1) invest in AI research and development (R&D); (2) unleash AI resources; (3) remove barriers to AI innovation; (4) train an AI-ready workforce; and (5) promote an international environment that is supportive of American AI innovation and its responsible use.
May, 2018	Sweden	AI Strategy	The strategy document serves as a reference to help the government to outline forthcoming policy initiatives aiming at strengthening Sweden's welfare and competitiveness by fully exploiting the benefits of AI. To this purpose, the Swedish strategy proposes to focus on the following priority areas: Education and training; Research; Innovation and use; Framework and infrastructure.
June, 2018	India	National Strategy for AI (#AIforAll)	The strategy aims to (1) enhance and empower Indians with the skills to find quality jobs; (2) invest in research and sectors that can maximize economic growth and social impact; and (3) scale Indian-made AI solutions to the rest of the developing world.
June, 2018	Mexico	Towards an AI Strategy	Priority areas: Governance, government and public services; R&D; Capacity, skills and education; Data infrastructure; Ethics and regulation.
July, 2018	Germany	Outlines of the goals of AI Strategy	To strengthen and expand German and European research in AI and focus on the transfer of research results to the private sector and the creation of AI applications.
November, 2018	Germany	AI Strategy (AI Made in Germany)	It focuses on 12 fields of action and 14 goals.
Fall, 2018	EU	AI Strategy	This analysis looks at the current state of development of AI technologies and at the policy mix and industrial policy tools that Europe and its Member States have to put in place to ensure that Europe is in the most advanced position in terms of developing applications of AI in industry.
December, 2018	Australia	Digital Economy Strategy	Australia's Tech Future details how Australia can maximise the opportunities of technological change by focusing on 4 key areas: 1/ developing Australia's digital skills and leaving no one behind; 2/ how government can better deliver digital services; 3/ building infrastructure and providing secure access to high-quality data; 4/ maintaining our cyber security and reviewing our regulatory systems.
December, 2018	Finland	Policy report on Ethical information policy in an age of AI	It outlines principles for fair data governance, including guidelines for the use of information and ethical values.
June, 2019	USA	The National AI Research and Development Strategic Plan: 2019 update	In this Strategic Plan, eight strategic priorities have been identified. The first seven strategies continue from the 2016 Plan, reflecting the reaffirmation of the importance of these strategies by multiple respondents from the public and government, with no calls to remove any of the strategies. The eighth strategy is new and focuses on the increasing importance of effective partnerships between the Federal Government and academia, industry, other non-Federal entities, and international allies to generate technological breakthroughs in AI and to rapidly transition those breakthroughs into capabilities.

Source: Author's work



and exist in the expanding digital world, managers must consider the following key dimensions of business ecosystems when making strategic decisions (Panetta, 2018). The first key dimension is the ecosystem strategy. Every company exists in many business ecosystems. They can be considered as dynamic networks of entities that interact with each other to create and exchange sustainable value for the participants in the respective ecosystem. Nowadays, the top technology investments of benchmark companies in digital ecosystems are BI and business analytics, followed by cloud services / solutions and digital marketing. One of the major problems is that progressive technology is difficult to implement on an industrial scale, and is usually created by start-ups. In this way, companies rely too much on 'external' innovation and/or do not create transformational technology. With such a strategic orientation, companies must plan to aggressively expand the partner pool or risk not being able to reach the so-called digital escape velocity.

Second, the degree of openness in the ecosystem is determined by the strategies, common goals and shared interests of the participants. The ecosystem can be public, private or hybrid. Many companies are de facto involved in a hybrid formation between a private and a public ecosystem. The degree of openness of the ecosystem has two main applications. The degree of change depends on the likelihood of new entrants and the breakdown of relationships and value. It defines the nature of relationships in ecosystems and how they are formed and maintained, as well as the nature of collaboration and competition.

Third, the level of involvement / inclusion of diverse actors in ecosystems. As connectivity increases, companies need to find a way to integrate modern IT technologies such as smart advisors and artificial intelligence into their business systems. The top management of the companies must realize and understand that the diversity in the ecosystem and the roles that people, business entities and 'things' play will develop and change depending on the situations, i.e. the constant situational change will determine the way decisions are made

Fourth, the type of relationships in ecosystems. The interconnection of 7 billion people and more than 30 billion Internet-connected devices in 2020 poses a significant challenge to business ecosystems. Currently, most companies maintain their relationships in ecosystems through digital platforms, where participants with different goals and objectives are connected on a commission basis. The platform provides basic integration, various business applications and management services for participants, but the speed of technology development can radically change these relationships.

Fifth, the creation and exchange of value in ecosystems. In addition to the monetary exchange of value, ecosystems can dynamically maintain leverage of information, reputation, services and other non-monetary forms of value for customers.

*Table 4. Ecosystem forecast profile, 2025*

Ecosystem	Ecosystem Subcategory	Estimated Total Sales, USD Trillion
B2B retail marketplace	Logistics Corporate banking Clothing	8.3
Travel and hospitality	Restaurants Hotels	3.6
Mobility	Auto and gasoline sales	2.0
Education		0.6
Housing	Mortgages	5.0
Digital content	Telecom services	3.3
Health	Private and digital health	6.0
Public services	Recreation and culture	4.4
Wealth and protection	Mutual funds	1.1
Global corporate services	Transport-support activities	2.9
B2B services	Legal Accounting Management of companies	9.6
B2B industrial marketplace	Machinery and equipment	9.4

Source: Adapted by: IHS World Industry Service; Panorama by McKinsey; McKinsey analysis; Catlin et al. (2018:2)

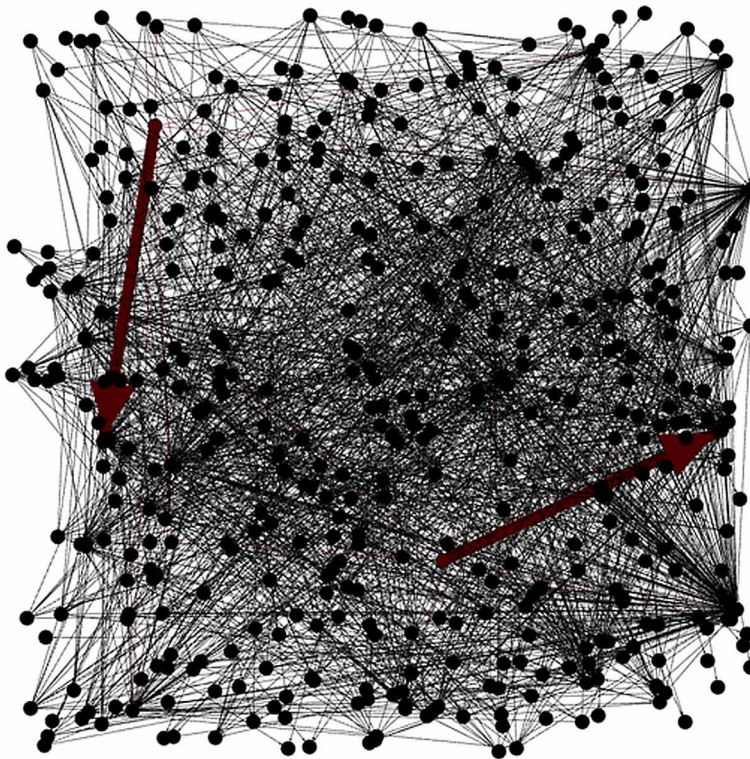
Sixth, diversity of industries. Ecosystem expansion can result in unexpected partnerships for organizations. Partners can be not only organizations from the primary industry, but also from similar and even remote industries (e.g. tourism and healthcare). Seventh, the complexity of multicomponent ecosystems. The probability of including large companies in multicomponent systems is high. It is crucial to understand how these systems interact, to identify potential discrepancies and overlaps, and to know their limitations and applications. It should be taken into account that some overlapping ecosystems may create a new ecosystem, while in other cases ecosystem overcrowding may occur. Eighth, technology. In order for companies to achieve market success, it is necessary to strategically integrate technology, information and business processes, which can be done by applying a business model that allows differentiation and generates value.

The complexity of the AI ecosystem is illustrated with the case of Italy. The dataset was downloaded from the website of the Agency for Digital Italy which is available at the following link: <https://ia.italia.it/en/ai-in-italy/>. The aim of this initiative is to map the Italian producers and users of AI solutions (startups, companies, public and private research entities, public administrations, etc.), to facilitate the

construction of relationships, the sharing of knowledge, and allow Italy to have the size of its strengths in the field of AI. The visualization covers 195 individual cases categorized by type of the enterprise<sup>4</sup>, industry sector<sup>5</sup>, town, region, and type of AI technology<sup>6</sup>. Visualizations of large graphs are useful to leverage the perceptual abilities of humans to find features in network structure and data. In this particular case, the open source software for graph and network analysis Gephi was used. Gephi uses a 3D render engine to display large networks in real-time and to speed up the exploration (Bastian, et al., 2009).

*Figure 1. Italian AI Ecosystem, visualization 1*

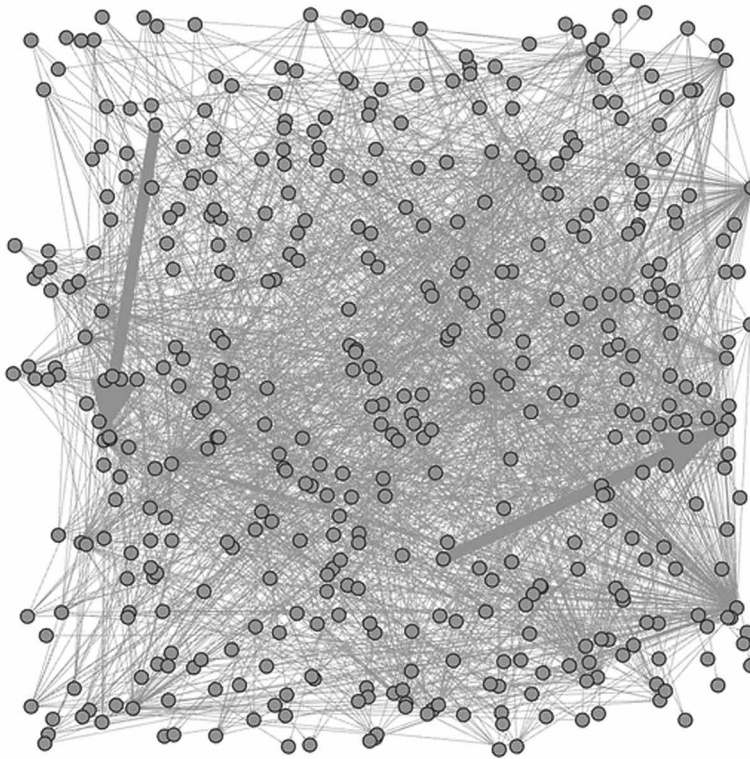
*Source: Author's work*



The visualizations of Italian AI ecosystem as a network (presented in Figure 1 and Figure 2) could be further analysed by adding filters which can select nodes or edges with thresholds, range and other properties available in the dataset. Dynamic

*Figure 2. Italian AI Ecosystem, visualization 2*

*Source: Author's work*



network visualization offer possibilities to understand structure transition within the network.

## **SOLUTIONS AND RECOMMENDATIONS**

The AI benefits are obvious but there are certain societal risks related to the diffusion of AI technologies in products and services which requires an open debate about AI governance. The main focus should be placed on developing an internationally recognised ethical and legal framework for the design, production and application of AI. This framework should be based on common AI principles and should provide a roadmap for protecting humanity by responsible uses of AI technologies. Unless AI is still at a relatively early stage of development and large scale industrial applications are yet to be developed, the societal challenges of AI applications should be explored and prioritized especially within the context of AI ecosystem.

## **FUTURE RESEARCH DIRECTIONS**

The safe and ethical use of AI research is a global challenge and there is still a significant policy and regulatory gap to fill in this respect (Delponte, 2018). Various entities participate in AI ecosystems which requires detailed risk-benefit analysis of their relationships and critical assessment of AI safety, transparency, and accountability within the ecosystem. The governance and management approach should not simply focus on the potential impact of AI on market performance and competitiveness but also on the social and ethical implications of an increased implementation of AI systems.

## **CONCLUSION**

Artificial Intelligence (AI) gains strong political support worldwide provided by the ambitious programs of governments and investments in AI-based technologies to reach global technology leadership. It is expected that AI will bring extensive effects on individuals, families, business and society as a whole. That is why, a combination of political, legal and technical factors must be implemented to govern responsible and transparent AI deployment on a global scale both by national governments and international organizations. There are several critical areas which should be addressed by the AI strategies and initiatives: those related to the inadequate internal technical capacity of SMEs and other organizations, those concerned with AI policy and regulatory risks, and those related to AI awareness level, penetration rate and level of business and social adoption of AI.

## **REFERENCES**

- Arthur, W. B. (2013). *Complexity economics: a different framework for economic thought*. SFI Working Paper: 2013-04-012, Santa Fe Institute.
- Arthur, W. B., Durlauf, S. N., & Lane, D. A. (1997). Introduction. In W. B. Arthur, S. N. Durlauf, & D. A. Lane (Eds.), *The Economy as an Evolving Complex System II* (pp. 1–14). Addison-Wesley.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*. Available at: <https://gephi.org/publications/gephi-bastian-feb09.pdf>

Beinhocker, E. D. (2006). *The Origin of Wealth – Evolution, Complexity, and the Radical Remaking of Economics*. Boston: Harvard Business School Press.

Catlin, T., Lorenz, J.-T., Nandan, J., Sharma, S., & Waschto, A. (2018). *Insurance beyond digital: The rise of ecosystems and platforms. Report*. McKinsey & Co.

Delponte, L. (2018). European Artificial Intelligence (AI) leadership, the path for an integrated vision. Study requested by the ITRE committee, Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies, PE 626.074- September 2018.

Gintis, H. (2006). *The economy as a complex adaptive system. A Review of Eric D. Beinhocker The Origins of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*.

Hamel, G. (2002). *Leading The Revolution*. Plume.

Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. (2013). *The Atlas of the Economic Complexity: Mapping Paths to Prosperity*. Massachusetts Institute of Technology and Center for International Development, Harvard University.

Montgomery, M. (1999). Complexity Theory: An Austrian Perspective. In D. Colander (Ed.), *Complexity Theory and the History of Economic Thought*. Routledge Press. <http://www.rasmusen.org/xpacioli/workpaps/99.06.Complexity.PDF>

Olley, D. (2020). *AI Report | Research Intelligence | Elsevier*. <https://www.elsevier.com/research-intelligence/resource-library/ai-report>

Panetta, K. (2018). *Ecosystems Drive Digital Growth*. <https://www.gartner.com/smarterwithgartner/ecosystems-drive-digital-growth/>

Parisi, G. (1999). Complex systems: A physicist's viewpoint. *Physica A*, 263(1), 557–564. doi:10.1016/S0378-4371(98)00524-X

Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J., Lyons, T., Etchemendy, J., Grosz, B., & Bauer, Z. (2018). *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University.

Singer, J. (2006). Framing brand management for marketing ecosystems. *The Journal of Business Strategy*, 27(5), 50–57. doi:10.1108/02756660610692716

## ENDNOTES

- <sup>1</sup> Gephi is an open-source software for visualizing and analysing large networks graphs. Gephi uses a 3D render engine to display graphs in real-time and speed up the exploration. You can use it to explore, analyse, spatialise, filter, clutterize, manipulate and export all types of graphs. <https://gephi.org/>
- <sup>2</sup> According to the results of KPMG research of 1400 senior corporate decision makers from 22 countries, representing seven primary business sectors.
- <sup>3</sup> Occasionally it is called the process-and-emergence perspective (Arthur, Durlauf, and Lane, 1997).
- <sup>4</sup> company, accelerator, research center, public administration, incubator, startup, university.
- <sup>5</sup> environment, agriculture, automotive, bioinformatics, economy and finance / fintech/insurance, education, power, precision industry, heavy industry, military, central public administration, territorial public administration (regional and local), advertising / marketing, applied research / statistics, robotics, health, security / cybersecurity, smart-cities, transportation.
- <sup>6</sup> Computer Vision, Natural Language Processing, Machine/Deep Learning, Recommender Systems, Robotics/Autonomous Systems, Expert Systems, Chatbot, Cognitive Cybersecurity.

## Section 2


# Responsible Application of AI Tools and Methods



# Chapter 5

## What We Should Have Learned From Cybersyn: An Epistemological View on the Socialist Approach of Cybersyn in Respective of Industry 4.0

**Dietmar Koering**

 <https://orcid.org/0000-0003-1390-428X>  
Arphenotype, Germany

### **ABSTRACT**

*Currently, a major topic is what changes will digitalization and the fourth industrial revolution bring to our society. It is clear that digital transformation of society and the introduction of new technologies will make many jobs obsolete. This process logically leads to the idea of a universal basic income (UBI). In this respect, the socialist project, Cybersyn, is of great interest because it constituted a prototype of a data- and people-related idea to solve this problem. The aim was to increase the country's production, while counteracting rising unemployment through a socialist paradigm, which is obviously pertinent to the development of Industry 4.0. Although Cybersyn can be considered as an early prototype and catalyst, today's exponentially greater computational power has made such systems real, and humans are often excluded from them. Human beings are also positively affected by digital transformation. Herein, the current work contributes to the ethical debate concerning the digital transformation of society.*

DOI: 10.4018/978-1-7998-4285-9.ch005

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

## INTRODUCTION

We live in a time where the digital transformation is causing a massive transformation of life as we know it (Stowasser, 2019). Technology has converted mankind into a real-time sensor that can measure almost everything and collect large amounts of data (Tegmark, 2017, pp. 23-31). This leads to a new evolutionary step, as promoted by Max Tegmark in *Life 3.0* (Tegmark, 2017). In fact, this would mean that data and humans would form a new symbiosis, in which the distinction between man and machine would be removed. In addition Yuval Harari declares in his book *Homo Deus*, that *Homo Sapiens* will be replaced by a new entity (Harari, 2017). If this constitutes a desirable future is not yet clear and is ultimately up to each individual to decide. However, this symbiosis, which has not yet occurred, is based on human-machine relations and its evolutionary process. This addresses the ethical problem of the relation of man and machine in the 20th Century.

An interesting difference in approaches of understanding the relation between humans and machines has been identified by Thomas Lamarre (Lamarre, 2012). Lamarre perceives two different types of relations of man to machine, one which refers to Martin Heidegger and the other to Norbert Wiener. Heidegger promotes a deconstructionist understanding of an “essence of technology” from a linguistic perspective, which considers everything through the lens of law (as a moral view) and being. Andreas Luckner writes that technical thinking and acting are therefore already contradictory forms of work, to the extent that they concern making use of available means to transcend labor (Luckner, 2008, p. 45). On the other hand, historically, the development of more efficient machines has aimed to increase commodification in order to make existing work more effective. Only the development of an Artificial General Intelligence (AGI) would finally succeed in the goal set by Andreas Luckner, i.e., to overcome the labors of work. Whether or not human beings would then achieve happiness constitutes a different question. Especially, the definition of happiness presents challenges, as different cultures probably possess different understandings of happiness, which leads to a general ethics problem in developing the goals for an AGI. For further exploration of Heidegger and his ideas about human-technical relations, please see Luckner (2008).

Norbert Wiener, however, employs a different approach. With his cybernetic model, Wiener explores the distinction between animal and machine. This, in the view of the French philosopher Gilbert Simondon, is dangerous as it reduces human beings and society to a machine (Lamarre, 2012). Simondon’s contention is probably related to a tactical mistake (Glanville, 2012) made by Norbert Wiener, who published the book *Cybernetics: Or Control and Communication in the Animal and the Machine* (Wiener, 1948) prior to his publication of *The Human Use of Human Beings* (Wiener, 1954). Ranulph Glanville assumes that if Norbert Wiener had

published his two books in reverse chronological order, cybernetics would be more appropriately valued as an ethical discipline (Fischer, 2019, p. 297). On the other hand, it is also possible that Wiener deliberately chose this order, as it was just at the end of the 2nd World War, and people may have had no time or will to engage in an ethical discussion of the man-machine relation. Today, we understand, however, that ethical implications are requisite, even in war time. Nevertheless, both views are currently important, and the boundaries between the differentiation are becoming blurred. Indeed, both views support understanding the relation of man and machine.

It is questionable, if this is a future for human beings, we should aim for. From the point of view of a digital utopist (Boguslaw, 1968), this enables a positive optimization of society and industry, if a consciousness about this transition exists to the human being. Norbert Wiener wrote in *Human Use of Human beings*, that “*We have modified our environment so radically that we must now modify ourselves in order to exist in this new environment. We can no longer live in the old one.*” (Wiener, 1954, p. 46). Of course, we might even ask why we have to modify ourselves? In today’s context it is not about the optimization of the body, far more it’s about the optimization of mind, our consciousness. Adaption itself, of course, is a central human ability, and also occurs in our interactions with software. According to Max Tegmark, human adaption can only take place by developing knowledge, which is expressed in the theory of Life 3.0 (Tegmark, 2017). Thomas Fischer expresses the process in another way, stating: “*In our efforts to maintain our well-being, we adapt to given circumstances and adapt our circumstances to our needs.*” (Fischer, 2019, p. 281). Fischer also writes that, from the perspective of cybernetics, human beings as a whole form a closed loop. Specifically, our environments, and specifically our living conditions, make adjustments within us and we make adjustments to our environments. We need to be aware about the changes and possibilities enabled by the digital transformation. Hence we focus on new jobs, which won’t be replaced by the digitalization or assists this process, to have a certain job guarantee. To master the complexity of new data, interactive and responsive environments, as well as new academic courses, to create knowledge for this complexity especially in urban environments are required (Koering, 2019). The remaining questions are: Who will steer this environment? Will it be a human or artificial general intelligence (AGI)? Is fair participation with AGI even possible? This addresses quite clear the ethical problem, why we as humans have to adapt to our modified environments. Stephan Kaufmann very critically assumes in his text “Digitization, class struggle, revolution” that digitization and Industry 4.0 are not unproblematic constraints, but a project of those who refer to themselves. It depends on these interests, what becomes reality and what remains only reverie. On the contrary, the workers and wage earners in the Industry 4.0 scenarios are scheduled as dependent variables.

They have to adapt to the “change”. They live in the passive: their leisure time and their work are digitized (Kaufmann, 2016, p. 2).

One early interactive and responsive environment was Cybersyn and/ with Cyberfolk in the 1970's, with the Opsroom (Figure 1). Even if the project could not predict its political end and thus went under, due to the military coup, which resulted in the dead of Allende and Cybersyn, the basic ideas from today's perspective are of enormous interest. Long before Industry 4.0, Cybersyn was an attempt to obtain information directly from the production front using state-of-the-art communication methods (Borchers, 2018, p. 77). However important is, that the decisions were still made by human beings. The operation room had seven chairs in total, which relates to a paper by George A. Miller with the title “The Magical Number Seven” (Miller, 1955). Essentially, it asserts that humans are in the position to remember seven plus or minus two “chunks” of information in short-term working memory. Thus, Stafford Beer chose seven seats to manage the incoming data and to make decisions in a participative way (Bonsiepe, personal communication, 2016). Things are different nowadays and the question is, where is the human being? In an interesting interview, Andreas Syska points out, that Industry 4.0 has been bypassed by human beings (Syska, 2015). Quite often decision are made by algorithms, which were of course programmed by human beings, but do not reflect human beings in the decision process, which is very effective, but on the other side creates a huge ethical problem. It leads to the fear that drives the digital transformation of the economy to job losses is real, especially as the debate over the universal basic income (UBI) is pursued. There is a risk that the UBI will become an industrial donation degenerating to combat poverty (Schwartz, 2010, p. 93). If there are no real opportunities offered, what human beings can acquire with an UBI, it won't help the society. Far more we can assume that also an UBI will result in higher rents, living costs, and so forth. Hence a political action is required. Eden Medina asserts about Cybersyn that “*It was a system designed to help the state regulate the nationalized economy and raise production without unemployment.*” (Medina, 2011, p. 211). This is indeed a very interesting quote, and is probably linked to the socialist ideas Allende, as it would be premature to talk about the effectiveness of Cybersyn in reducing unemployment, since the system has hardly been used (Bonsiepe, personal communication, 2019). Unemployment was a socialist thought desired by the project but not related to any specific policy (Espejo, personal communication, 2019). However, raising awareness in a timely manner can eliminate the risk of unemployment by making a career choice in advance and seeing what is needed in the future. Nevertheless, Industry 4.0 will also offer new jobs, perhaps even jobs that we have not previously considered. Other jobs will disappear in the near future. Additional details can be found in an interesting study: “*The Future of Employment: How susceptible are jobs*

to computerisation?” (Frey & Osborne, 2013). This is due to the fact that robotics, new ways of production, and AI will replace workers.

## **INDUSTRY 4.0**

Industry 4.0 or the fourth industrial revolution (Schwaab, 2017) is becoming a common term. It is interesting that the term revolution is used, as a revolution is an uprising of the masses, which results in a regime change. As the fourth industrial revolution is led by big global players, we can state, that this was not initiated by the masses. This is again in line with that, what Stephan Kaufmann, also with a socialist attitude, expressed. Maybe “Coup” would be an appropriate term, as it is a regime change, without considering the masses, the human beings. Nevertheless “coup” has a negative association. Quite interesting to this is, that Google initiated an AI ethics board in 2019, which was dissolved just one week after forming it (Levin, 2019). But it has to be said, that the ethics board was dissolved as the employees called for the removal due to the composition of the advisory board. Stafford Beer has shown this already in an interesting graphic (Figure 2). Figure 2 illustrates the homeostatic relationship of data and their exploited sources. It is critical to understand that exploitation in computer science simply means gathering information. Indeed: “*Exploitation is often negatively termed...meaning in computer science: Gathering Information.*” (Christian & Griffiths, 2016, p. 32). Hence it is about the exploitation of the masses by the well organized global players, which results in a class war.

Essentially for a well-organized Industry 4.0 is, that all becomes connected in real-time; everything can be optimized and traced. Clearly, commodification also constituted an integral part of the industrial revolution, in which the steam engine largely replaced human labor (Brynjolfsson & McAfee, 2014, pp. 6-7). Brynjolfsson and McAfee extend this, however, to claim that it is even more important to have machines that can complete cognitive tasks, and not solely physical ones (Brynjolfsson & McAfee, 2014, p. 91). This can be observed in the current development of Smart Cities and Industry 4.0, which are enabled by sensors that deliver massive real-time data from their environments to augment productivity. We might then ask, when the steam engine was replacing the muscle power, and later the computers the brain power, what is then replaced by the Industry 4.0?

The increases of the productivity with ideally fewer resources is only possible with Industry 4.0, which links to the concept of the IoT and Industry 4.0 (Nascimento Marques Junior, 2018) or the “Industrial Internet” (Evans & Annunziata, 2012), which has its roots in computerized manufacturing processes, or maybe even in Cybersyn (Clancey, 2017). It describes conscious production methods enabled

through networks, e.g., the “smart” factory. But while we talk about ethics, we need to admit that ethics is something, which needs a brain, a brain has intelligence and therewith consciousness. But we do not know, how consciousness is generated or can be expressed in an algorithm, or when AGI or Singularity will happen. Hence how is it possible then to speak about conscious production methods, if we not link human beings as main driver to this, something what has been already done in Cybersyn?

The nine pillars for the fourth industrial revolution are Big Data, Autonomous Robots, Simulation, System Integration, Internet of things, Cloud Computing, Additive Manufacturing, Augmented Reality and Cyber Security (Erboz, 2017). These pillars create a new complexity, which need to be managed in control-rooms, as these control-rooms somehow offer the possibilities of control and steering processes by human beings, as humans should be the most important part to these systems; maybe the human being presents the architrave, which is resting on theses pillars.

## **CYBERSYN**

The Project Cybersyn (Figure 3) was one of the first, if not the first, digital decision support system. Furthermore, it was a rare project that combined social responsibility, novel technologies, and design (Bonsiepe, 2009, pp. 35-62) realized in 1970 to 1973. Theoretically, the project was far more advanced than the resources of a peripheral country allowed at that time, especially in a period of political confrontation, when the government’s political project was directed against geopolitical hegemonic interests, which then regained their upper hand as part of a military coup. Also important in the debate over the theoretical background of Cybersyn was the thesis of the physician and neurologist Ross Ashby about the requisite variety, which states that a system to survive must produce a larger variety than the environment from which disturbances emanate (Bonsiepe, personal communication, 2016). It can be seen as a version of a “proto-internet”, an interactive and responsive environment. The common English name was Cybersyn, while the Spanish title of the project was SYNCO, as Cybersyn was not euphonic in the Spanish language. Actually, Cybersyn referred to a computer system that was connected to a network of telex and radio connections, termed “Cybernet” (Pias, 2007). Cybernet, was supposed to transmit daily the production figures from the nationalized companies in the six main sectors of Chile (energy, steel, copper, petrochemical, fishing and transport) to the headquarters. The data should be fed to the Cyberstride software and calculated in simulations to detect delivery bottlenecks and other issues early on. Another module called Cyberfolk was planned (Espejo, 2017, p. 43), where the Chileans, workers and employees report on terminals with their own happiness in real-time with the government (Medina, 2011, p. 89), production and distribution ideas (Borchers, 2018,

p. 77). Cyberfolk ideas were tested as part of INTEC's (the Institute for Scientific and Technological Development) contribution to the Project Cybersyn. Stafford's son, Simon Beer, designed the electronics of the Algedonometer. A group of the Cybersyn team had a weekly meeting to discuss longer term developments of the Cyberfolk project. Raul Espejo asserts, that: *"These electronics were tested in a couple of these meetings, one I remember well with Heinz von Foerster, in which us, the participants had a small device with a knob in our laps which we could individually move in a positive or negative direction, to express our views about the progress with the speaker's talk."* (Espejo, personal communication, 2019). The idea was, that the speaker could see the integration of the individual 'votes' (positions) in a device in the front of the room, which naturally was seen by the team, the observer and by the speaker at the same time. This was the most concrete design the team had at that time of an Algedonometer. Nevertheless, the idea of Cyberfolk to include peoples insights in real-time, shows again, that the people are bypassed in the fourth industrial revolution; human being became an accessories. This shows, how advanced the idea of an democratic governance already was. If we compare this to the latest judgment by the European Court that companies must systematically record working hours (Feldforth, 2019), it is obvious, that we need an ethical discourse about the inclusion of the people. However, this judgement links directly to Frederick Winsor Taylor, the founder of ergonomics. The aim of Taylor was to identify the "best way" of performing a work step in order to make the work measurable and to ensure machine-like, error-free, and efficient execution (Hessler, 2014). Human labor became, though Taylor's work, measurable and therewith humans could be "optimized" and perhaps even replaced by the usage of machines. Martina Hessler further notes that knowledge was transcribed and thus made into an objectified form of the individual body, whereby the workers were, to a certain extent, living machines. Hence, the idea was born already with Cyberfolk, but with a human friendly attitude, as it was about the democratic inclusion of the people, which was in addition also important for Allende (Espejo, 2017, p. 42).

The project Cybersyn was initiated under the Chilean President, Salvador Allende, with an socialist ideology<sup>1</sup>. It was managed by Fernando Flores, who was the political director, Raul Espejo as technical director and the British cybernetician Stafford Beer, who was the scientific director. The main discussion between Beer and Allende was about how such a cybernetic system would enable control. For Beer it was about the possibility which would enable Allende to take decisions, while for Allende this higher "stage" of the system was clearly for the people, as it was implicit to the socialist ideology of Allende. Having Stafford Beer as a lead scientist on the project probably relates to his book "Decision and Control – The meaning of Operational Research and Management Cybernetics" published in 1966 (Beer, 1966). In this book, Beer describes cybernetic systems to analyze

inventory systems in order to reach optimal decisions and to handle large complex and hazardous situations that might arise in industry, government, and business environments. Hence the idea was, to manage and steer the production, somehow close to the idea of Industry 4.0, but decades earlier. These keywords for a project in 1973 demonstrates further the actual link to today's Big Data debate and IoT. It is obvious that the digital transformation will change the environment - and there is a well-founded fear of high unemployment. Industry 4.0 does not contradict this, it is more about creating a basic income to overcome this problem. Hence we need to discuss the role of human being in this environment. The role of being in interactive and responsive environments alongside algorithms.

Cybersyn ended after three years on 11 September 1973, as Allende's regime was overthrown. The new installed dictator Pinochet did not need real-time centralized planning or to monitor the moods of citizens. However, Beer had already noted the importance of this project, stating that: "...*information is a national resource*" (Beer in Morozov, 2014). With this estimation, Beer was ahead of his time, examining the actual context and ownership of data, as well as the open data policies of certain cities. Here, we have to agree with Morozov regarding his question about the means of data production that cannot be reduced to its technological dimensions.<sup>2</sup> Georg Jochum frames it quite nicely, by asserting that "*In view of the further development of cybernetic technologies, his (Stafford Beer) project of emancipatory cybernetics could not only be feasible, but even made necessary to prevent the triumph of cybernetic capitalism which comes along with the digital despotism.*" (Jochum, 2017, p. 543).

## CONCLUSION

While Cybersyn was designed to regulate the economy and raise production in accordance with socialist ideology, Industry 4.0 increases production in accordance with a capitalist ideology. The key aim of Industry 4.0 is to free human beings from processes in which computers can make faster and better decisions in production and elsewhere. Of course, this new situation automatically results in massive unemployment. Interestingly, Cybersyn held quite the opposite ideology, in which human beings constituted the central part of Cybersyn. Specifically, decisions were made by people in a participative manner, and not by algorithms. The problem of rising unemployment cannot be effectively addressed by Cybersyn, as the project ended relatively quickly. In addition, in the socialist approach, a core goal is the prevention of unemployment. Currently, the issue of widespread unemployment due to digitalization is addressed with UBI. Consequently, Cybersyn, Cyberfolk, and the democratic inclusion of people is instructive, since a basic income does not solve the problem, but only constitutes a transitional state. In fact, interactive



and responsive environments are requisite, which are steered by human beings, to understand the complexity of today's manufacturing practices under the conditions of Industry 4.0 in terms of the impact on our environments and the way that we, as human beings, participate in an artificial intelligence (AI)-driven transition. A fear recently exists that human beings are bypassed by this transition, and we live along-side algorithms. In other words, human beings have lost control over their lives and critical societal processes. The issue of what kind of control can be enabled by Cybersyn forms the primary discussion between Allende and Beer. This also constitutes an ethically significant debate, as digital transformation will affect our everyday lives and environments. Interactive and responsive environments are certainly necessary in order to receive qualitative benefits from digitization, but humans must be the objects of focus, and not only what AI enables humans to accomplish. Especially, scientists are tasked to develop and implement a universal code of ethics. Although this issue presents a wicked problem, a great need exists to begin a fruitful debate.

## REFERENCES

- Beer, S. (1966). *Decision and Control – The meaning of Operational Research and Management Cybernetics*. John Wiley & Sons.
- Boguslaw, R. (1968). *New Utopians: Study of System Design and Social Change*. Spectrum Books.
- Bonsiepe, G. (2009). *Entwurfskultur und Gesellschaft: Gestaltung zwischen Zentrum und Peripherie* (1st ed.). Birkhaeuser. doi:10.1007/978-3-0346-0389-8
- Borchers, D. (2018). Das Cybersyn-Projekt Wie Chile einst die Zukunft der Planwirtschaft entwarf. *c't Retro 2018*, 77. Retrieved from <https://www.heise.de:https://www.heise.de/select/ct/2018/27/1541215368236612>
- Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a time of Brilliant Technologies*. Norton & Company.
- Christian, B., & Griffiths, T. (2016). *Algorithms to Live by - The Computer Science of Human Decisions*. Henry Holt and Company LLC.
- Clancey, R. (2017). *Here Lies Project Cybersyn: Salvador Allende and Stafford Beer's Cybernetic System of Coordination for Chile's Economy (1971-1973)*. Strata.
- Erboz, G. (2017). How To Define Industry 4.0: Main Pillars Of Industry 4.0. *Conference: 7th International Conference on Management (ICoM 2017)*.

Espejo, R. (2017). Cybernetic Argument for Democratic Governance: Cybersyn and Cyberfolk. In L. C. Werner (Ed.), *Con-Versations Vol.1 cybernetics: state of the art* (pp. 34-57). Berlin: Universitätsverlag der TU Berlin.

Evans P. C. Annunziata M. (2012). Retrieved from [https://www.ge.com/docs/chapters/Industrial\\_Internet.pdf](https://www.ge.com/docs/chapters/Industrial_Internet.pdf)

Feldforth, O. (2019). *Arbeitszeit klar erfassen - aber wie?* Retrieved from <https://www.tagesschau.de/>: <https://www.tagesschau.de/wirtschaft/eugh-arbeitszeiten-107.html>

Fischer, T. (2019). Kybernetik. In T. Schoeler, S. Hoeltgen, & J. F. Maibaum (Eds.), *Medientechnisches Wissen* (pp. 275–301). De Gruyter.

Frey, C. B., & Osborne, M. A. (2013). *The Future Employment: How susceptible are jobs to computerisation?* Retrieved from <https://www.oxfordmartin.ox.ac.uk/>: [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf)

Glanville, R. (2012). Radical constructivism = second-order Cybernetics. *Cybernetics & Human Knowing*, 4(4), 27–42.

Griffiths, T., & Christian, B. (2016). *Algorithms to Live By: The Computer Science of Human Decisions*. Henry Holt and Co.

Harari, N. Y. (2017). *Homo Deus - A Brief history of Tomorrow*. Vintage / Penguin Random House. doi:10.17104/9783406704024

Hessler, N. (2014). Die Halle 54 bei Volkswagen und die Grenzen der Automatisierung. *Zeithistorische Forschungen/Studies in Contemporary History*, 11, 56-76. Retrieved from <https://zeithistorische-forschungen.de/1-2014/id%3D4996>

Jochum, G. (2017). *Plus Ultra« oder die Erfindung der Moderne: Zur neuzeitlichen Entgrenzung ...* Bielefeldt: transcript.

Kaufmann, S. (2016). *Digitalisierung, Klassenkampf, Revolution*. Retrieved from <https://www.rosalux.de/>: [https://www.rosalux.de/fileadmin/images/publikationen/Analysen/Analysen33\\_Digitalisierung.pdf](https://www.rosalux.de/fileadmin/images/publikationen/Analysen/Analysen33_Digitalisierung.pdf)

Koering, D. (2019). Conscious City Laboratory - Explorations in the history of computation, cybernetics, and architecture: Foresight for artificial intelligence and human participation within cities. Universitätsverlag der TU Berlin. DOI: 10.14279/depositonce-8466

Lamarre, T. (2012). Humans and Machines. *Inflexions*, 5, 29–67.

- Levin, S. (2019). *Google scraps AI ethics council after backlash: Back to the drawing board*. <https://www.theguardian.com/technology/2019/apr/04/google-ai-ethics-council-backlash>
- Luckner, A. (2008). *Heidegger und das Denken der Technik*. Bielefeldt: Transcript.
- Medina, E. (2011). *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. MIT Press. doi:10.7551/mitpress/8417.001.0001
- Miller, A. G. (1955). The Magical Number Seven, Plus or Minus Two Some Limits on Our Capacity for Processing Information. *Psychological Review*, 101. Retrieved from Psychological Review: <https://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf>
- Morozov, E. (2014). *The Planning Machine: Project Cybersyn and the origins of the Big Data nation*. The New Yorker.
- Nascimento Marques, M. R., Jr. (2018). Embedded Agent based on Cyber Physical Systems: Architecture, Hardware Definition and Application in Industry 4.0 Context. In *15th International Conference on Informatics in Control, Automation and Robotics* (pp. 584-591). Retrieved from Center of Computational Sciences, Federal University of Rio Grande, Rio Grande, Brazil: [www.researchgate.net](http://www.researchgate.net)
- Pias, C. (2007). *Defense of Cybernetics. A Reminiscence*. WEB.
- Schwaab, K. (2017). *The Fourth Industrial Revolution*. Crown Business.
- Schwartz, E. M. (2010, May). *Poverty reduction for profit? A critical assessment of the Bottom-of-the-Pyramid Approach and of the 'Opportunities for the Majority'-Initiative of the Inter-American Development Bank*. Retrieved from University Vienna: <https://core.ac.uk/download/pdf/11590712.pdf>
- Stowasser, S. (2019). *KI verändert die Welt – auch die Arbeit*. Retrieved from <https://www.wissenschaftsjahr.de/2019/neues-aus-der-wissenschaft/das-sagt-die-wissenschaft/ki-veraendert-die-welt-auch-die-arbeit/>
- Sweeting, B. (2019, April 1). Applying ethics to itself: Recursive ethical questioning in architecture and second-order cybernetics. *Kybernetes*, 48(4), 805–815. Advance online publication. doi:10.1108/K-12-2017-0471
- Tegmark, M. (2017). *Life 3.0 - Being Human in the Age of Artificial Intelligence*. Allan Lane.
- Tegmark, M. (2017). *Life 3.0 - Being human in the age of Artificial Intelligence*. Penguin Random House UK.

### ***What We Should Have Learned From Cybersyn***

Wiener, N. (1948). *Cybernetics: Or control and Communication in the Animal and the Machine*. MIT Press.

Wiener, N. (1954). *The Human use of Human Beings. Cybernetics And Society*. Doubleday Anchor Books.

## **ENDNOTES**

- <sup>1</sup> Which of course contradicts the ideology of Industry 4.0, and maybe it might not be even fair, to compare Cybersyn to Industry 4.0, as they have different roots.
- <sup>2</sup> Information regarding SYNCO and its theoretical foundations relates to personal communication with the interface designer, Gui Bonsiepe, in March 2016 and onwards. Also relevant is the essay by Gui Bonsiepe, “Der Opsroom – zum Eigensinn der Peripherie.”

# Chapter 6

## The Transformation and Enterprise Architecture Framework:

### The Applied Holistic Mathematical Model for Geopolitical Analysis (AHMM4GA)

**Antoine Trad**



<https://orcid.org/0000-0002-4199-6970>  
Independent Researcher, Croatia

#### ABSTRACT

*This chapter proposes the applied holistic mathematical model for geopolitical analysis (AHMM4GA) that is the result of research on societal, business/financial, and geopolitical transformations using applied mathematical models. This research is based on an authentic and proprietary mixed research method that is supported by an underlining mainly qualitative holistic reasoning model module that punctually calls to quantitate functions. The proposed AHMM4GA formalism, attempts to simulate functions to support empirical processes.*

#### INTRODUCTION

The AHMM4GA can be used to implement a Decision Making Systems for Geopolitical Analysis (DMS4GA) that can integrate in the country's, region's or organisation (or simply the *Entity*) analysis process. The AHMM4GA uses a Natural Language Programming (NLP) that can be easily adopted by the project's

DOI: 10.4018/978-1-7998-4285-9.ch006

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

development teams for implementing Geopolitical Analysis (GA) (Myers, Pane, & Ko, 2004; Kim & Kim, 1999; Della Croce & T'kindt, 2002), to model GA or complex Enterprise Architecture (EA) blueprints. The uniqueness of the author's Research & Development Project (RDP) is that the AHMM4GA promotes a holistic structure, for the alignment of various frameworks, standards and strategies to support transformation projects (Farhoomand, 2004). The proposed AHMM4GA has been applied to verify various cases in different fields and domains; and the results were more than satisfying; like for example, it in the cases: 1) financial services engineering and evaluation, to support the detection of financial irregularities and crimes; 2) to analyse the rise and the 1975 fall of the Lebanese business ecosystem; 3) to assess the risks of the Lebanese Islamic business strategy... (Trad, 2019d).; 4) to analyse gigantic financial manipulations like the ones, related to fraud and money laundering that damage many emerging countries, and this case it is related to the Universal Banking System (UBS) (Stupples, Sazonov, & Woolley, 2019), in which trillions of US dollars are hidden; in the moment where the world is suffering from recessions and other major problems; and 5) many other related cases.

GA or Transformation Projects (or simply the *Project*) are managed as a set of separate black-boxes, where their structures are a hairball; which is called the *Entity*, which owns an Information and Communication System (ICS). The DMS4GA is used to solve GA problems by offering a set of possible solutions in the form of explanations and GA recommendations; by using a central qualitative method based on a beam search (heuristic tree) that uses quantitative methods at its nodes. The proposed AHMM4GA's implementation is very complex and needs a profound understanding of many complex fields. The DMS4GA' actions produce solutions, which have the form of geopolitical, societal and managerial recommendations. A DMS4GA is a multi-objective, multi-project, multi Critical Success Factor (CSF). This chapter's background combines knowledge management, innovative DMS4GA approach, enterprise architecture, heuristics/mathematical models, ICS management, societal/geopolitical transformation initiatives and GA related fields (Goikoetxea, 2004; Tidd & Bessant, 2009). As shown in Figure 1, where the major strategic field trend is Artificial Intelligence (AI) based *Projects*; so the author concludes that building an innovative AHMM4GA is in the line of innovation (Cearley, Walker & Burke, 2016; Thomas, 2015; Ho, Xu & Dey, 2010). An AHMM4GA instance enables the implementation of a generic and cross-functional reasoning engine that is mainly based on: 1) mainly geopolitical CSF classification mechanisms; 2) an adapted qualitative heuristics tree research method; and 3) a set of quantitative modules that can be triggered from the tree's nodes. The author based his RDP uses existing industry standards, like for example The Open Group's Architecture Framework's (TOGAF) and its Architecture Development Method (ADM) (The Open Group, 2011a; Tidd & Bessant, 2009). The AHMM4GA is strategy driven and

Figure 1. Technology Trends  
(Cearley, Walker, & Burke, 2016).



is agnostic to any application fields. As shown in Figure 1, the author's is founded on DMS4GA that uses existing standards (Johnson & Onwuegbuzie, 2004).

Today, there are many mathematical models, technology standards and EA methodologies that can be used to implement a *Project* (Gartner, 2016), but all of them lack a systemic holistic approach. The GA manager or a transformation manager (or simply the *Manager*) can integrate an AHMM4GA based DMS4GA in the transformation roadmap to support its complex and risky analysis (Zaiane & Ben Moussa, 2018; Trad & Kalpić, 2017b, 2017c, 2018a, 2018b; Thomas, 2015; Tidd, 2006). The RDP's methodology is based on: 1) a multi-domain literature review; 2) a qualitative methodology; 3) a quantitative methodology; and 4) an engineering oriented Proof of Concept (PoC) (or a controlled experiment); which is the optimal methodology applied in complex analysis, information technology,

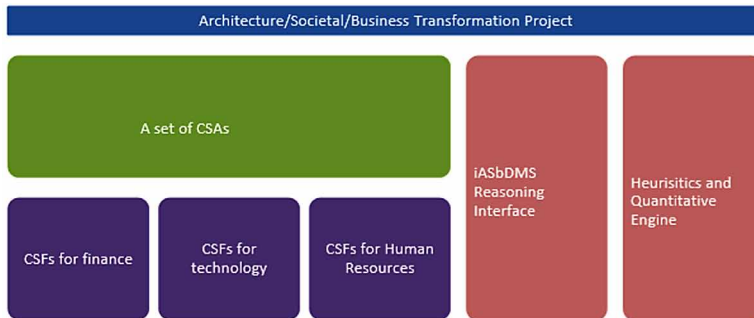
The Transformation and Enterprise Architecture Framework applied mathematics and other types of engineering projects (Easterbrook, Singer, Storey, & Damian, 2008). For a successful integration of the AHMM4GA in *Projects*, the *Manager's* profile and role are crucial and his or her (for simplicity, in further text – his) decisions are supported by the selection and implementation of CSF. A holistic system approach is the optimal choice to integrate an AHMM4GA in *Projects* (Daellenbach & McNickle, 2005; Trad & Kalpić, 2017d). The DMS4GA interacts with geopolitical analysts by means of an interface in order to manage the CSFs and to launch the reasoning process, as shown in Figure 2, where the Geopolitical module (Gm) interacts with other modules.

*Figure 2. The RDP's concept (Trad & Kalpić, 2017d).*





*Figure 3. Structure of critical success descriptors and heuristics based decision making.*



the main interaction is done by CSF sets and the reasoning engine. In *Projects*, there is a need to offer robust GA, because of the proactive need to avoid major problems.

## FINANCIAL TERROSRISM AND RELATED IRREGULARITIES

### Finance Technology and Cybercriminality

Finance and Technology (FinTech) is an essential factor for the stability of organisations, governments and countries, because of the following facts:

- FinTech aims to change traditional environments to become interactive and offer financial services using intelligent endpoints. Although FinTech can be used to tackle financial Cybercriminality, it seems that the countries that support massive financial misdeeds are making the largest investment in these innovative technologies (Ravanetti, 2016).
- Electronic money or (e)money, would make money more abstract and difficult to trace; individuals and institutions in countries that have the culture of financial secrecy, national destabilizations and arbitral confiscation.
- Blockchain technologies will dominate the leading financial giants and will cause the domination of (e)money, this new media would lead to disappearance of leading currencies.
- Blockchains and (e)money is a framework that supports cryptocurrency that supports exchange of currencies in a digital encryption form.

## **A Radical Transformation Process**

Microartefacts based automated finance, is a term that is (or can) be used for an ICS that supports partners, who don't know each other to commit trustful financial operations and to share the logging records of all financial transactions, in a transparent manner. This sharing of logging records is virtual and is distributed to all partners on the internet who use their networks to execute transactions; which can be a security breach.

## **Automated Financial Processes**

An actual (e)money systems allow (e)transactions that include parties from different countries. (e)Signatures secure such (e)transactions and they provide an important part of the final financial solution, where the main benefits can be lost if a trusted third party is needed to prevent security breaches; and protects against UBS suspicious acts. The solution to the supplementary problem is to use an highly secure peer-to-peer ICS. Partners need to control (e)transaction's execution using automated mechanisms; which are capable of detecting gigantic financial misdeeds (Stupples, Sazonov, & Woolley, 2019).

## **The Automated Finance Critical Success Factors**

Based on the literature review and associated evaluation processes, the CSFs are evaluated and are explained below.

## **THE RESEARCH AND DEVELOPMENT PROJECT'S STRUCTURE**

### **The RDP's Cluster and Questions**

The main topic of the RDP is related to *Projects* and their risk management capacities. The ultimate global RDP's Research Question (RQ) is: "Which transformation manager characteristics and which type of support should be assured in the implementation phase of a transformation project?". This long RDP has many phases. Where in this chapter's or phase's RQ is: "Can an applied holistic mathematical model support geopolitical analysis in the context of global financial instability?".

## **The Research's Uniqueness**

The uniqueness and the lead of the author's works, can be verified by searching with the scholar engine, within Google's online search portal, in which the author combined the previously mentioned keywords and key topics; the results show very clearly the uniqueness and the absolute lead of the author's works/framework, methodology, research and recommendations in the mentioned scientific fields and that can be considered as an important jumpstart for the future industrial use (Trad, 2019e). From this point of view and facts, the author considers his consequent long-life works on the mentioned topics as unique, innovative, credible and ultimately useful; especially for the support the chapter's topic.

## **The Research Gap**

The author would like to contribute to enhance the success rate of *Projects*. To close or at least narrow the gap in the mentioned research field, the author proposes a holistic approach that unifies the following:

- The AHMM4GA that maps to all the *Project's* resources and components.
- The CSA and CSF management.

## **The Research Basics**

The AHMM4GA supports GA and EA risk concepts (Busenitz, 2014) and is supported by an Knowledge Management System for GA (KMS4GA). The actual chapter and the PoC are also a part of the Selection management, Architecture-modelling, Control-monitoring, Decision making, Training management, Project management, Finance management, Geopolitical management, Knowledge management and Implementation management Framework (SmAmCmDmTmPmFmGmKmImF, for simplification reasons, in further text the term the Transformation Research & Architecture Development framework (*TRADf* will be used). *TRADf* is composed of the following modules:

- “Sm”: for the selection management of the Framework.
- “Am”: for the architecture and modelling strategy that can be applied by the Framework.
- “Cm” for the control and monitoring strategy that can be applied by the Framework.
- “Dm” for the decision-making strategy that can be applied by the Framework.
- “Tm” for the training management of the Framework.

### ***The Transformation and Enterprise Architecture Framework***

- “Pm” for the project management strategy that can be applied by the Framework.
- “Fm” for the financial management’s support to the Framework.
- “Gm” for the Geopolitical mind-mining to the Framework, which is this chapter’s focus.
- “F” for Framework

The *TRADf* is not a black-box product to be applied as-is, it is rather a transformation strategy that offers recommendations and vision on how to implement a GA. Where here the main focus is Gm.

### **The Mathematical Model’s Basics**

Polderman and Willems (Polderman & Willems, 1998) argue that a Mathematical Model (MM) is a subset of real world’s behaviours and possibilities; and that a MM is a description of a limited reality. In this RDP, the reality is a GA case. Once a MM is implemented, it can offer a certain subset of possible analysis artefacts, solutions or explanations. The MM basics, are:

- An MM should provide abstractions of a real world of a physical system (Hinkelmann, 2016).
- Modeling technics are a descriptive design process, which validates MM principles (Sankaralingam, Ferris, Nowatzki, Estan, Wood, & Vaish, 2013).
- The usage of EA, can be used by an MM.
- The gap between the MM based GA’s adoption and its usage is still huge today (Syynimaa, 2015).
- An MM that is optimized for GA that uses CSFs (Dogan, Çalgici, Arditi, & Gunaydin, 2015).
- A generic MM, like the proposed AHMM4GA (Giachetti, 2012; Kim & Kim, 1999).
- An applied MM is the description of an *Entity* using mathematical concepts and language (Sankaralingam, Ferris, Nowatzki, Estan, Wood, & Vaish, 2013).
- Multi-criteria or a multi-factor model for decision making needs a mixed method based on qualitative and quantitative criteria (Zandia & Tavana, 2011; Kim & Kim, 1999).
- An MM is optimal for an empirical engineering research project (Easterbrook, Singer, Storey, & Damian, 2008).

The CSF-based RDP would use the AI based GA in *TRADf's* Gm. The RDP's phase 1 (represented in automated tables), which forms the empirical part of the RDP, checks the following CSAs:

- RDP, is synthesized in Table 1.
- AHMM4GA as a structure, is synthesized in Table 2.
- Applied Case Study (ACS) for the PoC, is synthesized in Table 3.
- ICS, is synthesized in Table 4.
- ADM, is synthesized in Table 5.
- Finance Engineering (FE) is in table 6.
- Human Resources (HR) is in table 7.
- KMS4GA, synthesized is in Table 8.
- DMS4GA, synthesized is in Table 9.
- GA case, synthesized in Table 10.

The final, Table 11, aggregates tables 1 to 10. The GA is supported mainly by a mixed ACS from two case studies: 1) insurance domain (Jonkers, Band, & Quartel, 2012a), for purely technical ACs; and 2) a geopolitical case related to Lebanon (OECD), for GA's applications, as shown in Figure 4.

Figure 4. The RDP's construct.

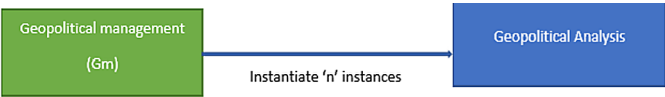
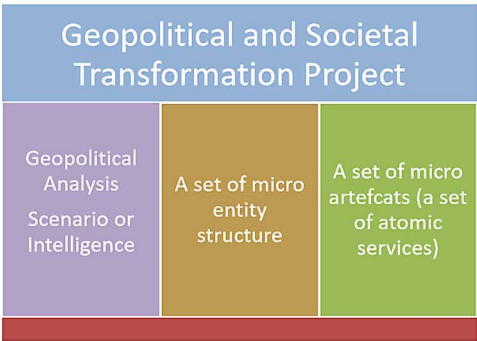


Figure 5. The RDP's construct.



The *Entity's Manager* geopolitical strategic decisions, should be made in a just-in-time manner, by using data/information from various credible sources. A framework proposes a strategy for governing an *Entity*, as shown Figure 5.

This chapter is a part of the RDP cluster that has produced many articles, literature reviews, usable items and microartefacts. In this chapter, like in the other author's works, the minimum of sections from previous articles are reused, for the better understanding of this complex concept and the GA capacities and that can be considered as a pioneering approach in the field. The GA is generic and is founded on a generic *TRADf* which in turn is based on the ADM (The Open Group, 2011a).

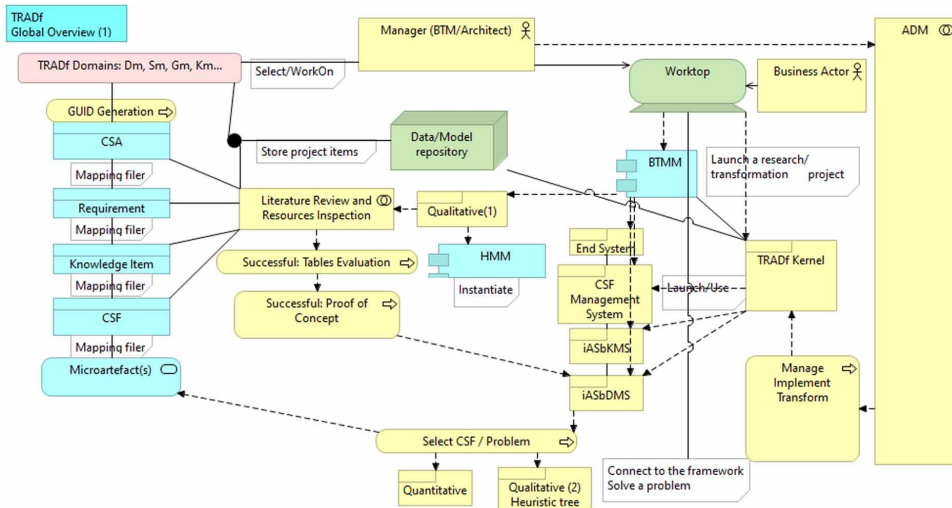
## **The Mechanics of CSFs Evaluation**

The RDP for GA starts with phase 1 (PHASE\_1) called the feasibility phase, which checks if the whole *Project* and its GA approach, is feasible. Then tries to evaluate the success rate using the most important CSFs, which are evaluated using the following rules:

- Rule 1: labelled the reference checker, all types of used references, should be credible and are estimated by the author; the notions of official ranking is less important and are ignored.
- Rule 2: labelled the change launcher, *Projects* like GA are the result of *Entity* or organisational changes in regions, the references are evaluated as presented in the previous point (Rule 1).
- Rule 3: labelled the logic checker, an applied modelling language or NLP should be used in a limited manner, in order to make the *Project's* GA manageable and not too complex.
- Rule 4: labelled the organisational construction, the ADM is considered to be mature, unfortunately that does not mean that *Projects'* GA are successful and in fact their success rate is very low.
- Rule 5: labelled the *Project* iteration management, the ADM is appropriate for any *Project's* GA iterative management and interface with *TRADf's* iterations.
- Rule 6: labelled the aggregation phase, if the aggregations of all the *Project* GA's CSA/CSF tables are positive and exceeds the defined minimum the *Project's* GA continues to execute the PoC and can be used for a problem solving.

As already mentioned, the six previous rules or steps are organised, parsed and weighted by the internal DMS4GA' engine, as shown in Figure 6.

Figure 6. The TRADf interactions.



## Research Work's Pattern and Concept Formulation

This chapter like all the other author's works have the same construct, which looks as follows (Trad & Kalpić, 2020a):

- An introductory part that explains the overall subject related to the phase's RQ.
- The RDP part that explains the research concept.
- The ACS and PoC related to the final experiment.
- The ICS, ADM, KMS4GA and DMS4GA parts, represent sections in the work's RQ specific context and integration.
- A specialized part, like in this cases of the GA contains a specific section and its literature review and synthesis.
- Each part (or CSA) contains a table of selected and weighted CSFs.
- The conclusion and recommendations that summarizes and concludes the research work.

## The RDP's CSFs

The GA's phase 1 processing, is based on CSA, CSF and KPIs evaluation (Trad & Kalpić, 2019e) and uses for example the following enumeration of CSFs: 1) Modelling; 2) Critical Success Factors; 3) References; 4) ArchitectureDevMethod; 5) Technologies; 6) Governance; 7) Transformation\_TRADf; and 8) Leading\_

TRADf offers a Mathematical Language (ML) that can be used to describe and implement a GA scenario (Türkmen, & Soyer, 2020; Goikoetxea, 2004), which can be used to solve *Project* problems.

	Critical Success Factors	KPIs	Weightings Ranges	Values
1	CSF_RDP_Modelling_GA_Cases	HighlyV easible	From 1 to 10.	10
2	CSF_RDP_CSF_Selection&Tuning	PossibleClassification	From 1 to 10.	10
3	CSF_RDP_References_Checking	AutomatedExists	From 1 to 10.	10
4	CSF_RDP_ADM_Integration	IntegrationPossible	From 1 to 10.	10
5	CSF_RDP_Technologies4GA	Advance dStage	From 1 to 10.	10
6	CSF_RDP_Governance_Deployment	Advanced	From 1 to 10.	9
7	CSF_RDP_Transformation_GA	IntegrationPossible	From 1 to 10.	8
8	CSF_RDP_Leading_TRAIDf	Possible	From 1 to 10.	9

Eval RDP

**RESULT:** 9,5

Entity **RDP** AMM ACS ICS ADM FIN HR KMS DMS GA P1

TRADf proposes a holistic approach to analyse events and eventually manage possible risks to help in avoiding major geopolitical pitfalls. Traditionally, complex risk concepts, were associated with a single origin or factor; 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 GA, for applies for example: 1) ethnicity; 2) finance (which is this chapter's focus; and 3) other factors...

Interconnecting divers fields, is enabled by the following evolution facts:



- The generalisation of the alphabet gave the possibility to communicate and prose complex problems. The Greeks and other civilization, inherited their alphabet from the Semite Phoenicians.
- The Semite Phoenicians introduced counting systems (numerical characters), geometry and arithmetic (Ball, 2010).
- Evolution of various mathematical fields like heuristics geometry, algebra... (Martin, 1981).
- The establishment of EA standards, that in turn have roots in Unified Modelling Language (UML).
- Business Process Management (BPM), Petri nets and Entity Relationship Management (ERM) are important ICS and EA domains where the usage of MM is fundamental to expose their abstraction... (Chu, & Xie, 1997).
- The AHMM4GA that inherits most of the previous evolutions and proposes a holistic approach to various domains.

## **MM's Generic Structure**

The MM's generic structure is used to resolve various types of GA scenarios. The use of a holistic methodology in the form of MM's generic structure can insure a successful outcome, or in the worst case, try to predict its pitfall through GA. A MM's generic structure represents the mapping relations between *Project's* resources, like CSFs, references, modules, microartefacts and *Entity's* resources that are not mutually exclusive. GA plan is generated by the DMS4GA' beam-search heuristics engine to realize enterprise transformation processes, using critical success areas and factors (Giachetti, 2012; Kim & Kim, 1999; Della Croce & T'kindt, 2002).

## **Critical Success Areas, Factors and Decision Making**

AHMM4GA is based on CSAs which are categories of sets of CSFs, where in turn, each CSF is a set of selected Key Performance Indicators (KPI), where: 1) each CSA corresponds to a *Project* domain, like for example, finance or political parties; 2) each CSF corresponds to a set of GA requirements, like for example, the *Entity's* accounting balance sheet finalization; and 3) each KPI corresponds to a single GA transformation or an *Entity* requirement; where the CSF/KPI elements interact with the ADM cycles. For each a prices *Project's* GA problem type, a DMS4GA qualified user can define the initial set of CSAs and CSFs. CSFs are important for the mapping between the GA requirements, microartefacts, *Entity's* organisational items to the AHMM4GA structure (Nilda Tri & Yusof, 2009; Peterson, 2011). The AHMM4GA's qualitative heuristic algorithms and punctual qualitative analysis can be used to evaluate for example the GA's special cases, in each CSA, where CSFs

can be internal or external; like: 1) the *Project's* gap analysis is an internal CSF; and 2) *Entity's* resources predictions is an external one as shown in Figure 4. Once the *Project's* GA initial set of CSFs have been selected, then the *Project's* members can use the AHMM4GA based DMS4GA to query for analysis results and possible solutions. The DMS4GA relates CSFs which map to a *Project's* GA requirement to a unit of work or microartefact (Trad & Kalpić, 2017b, 2017c).

## **The GA Model's Unit of Work**

A holistic alignment and classification of all the *Project's* resources for a GA execution, must be done, so that the unbundling and analysis processes can start. A holistic alignment needs also to define the AHMM4GA Unit of Work (UoW) or a basic microartefact. Using the "1:1" mapping concept, the microartefact is represented with a class diagram. Such a mapping concept is based on an automated naming convention that can link all the *Project's* resources for GA. The mapping concept supports the interoperability between all the *Project's* modules and enables the use of ML microartefacts that include the needed knowledge and intelligence support for GA (Mehra, Grundy, & Hosking, 2005; Scherer & Schapke, 2011).

## **AHMM4GA Microartefacts for GA**

An ML microartefact is any *Project* microartefact that is a part of the AHMM4GA and which interacts with a multitude of *Project* microartefacts in a coordinated manner to support a GA. An ML microartefact uses the ADM to assist the *Project's* GA implementation process (The Open Group, 2011a). The AHMM4GA includes various types of mechanisms that use heuristics scenarios to make the *Project's* GA integration more flexible and to avoid the classical and ridiculous archaic quantitative analysis and offer a holistic collaborative DMS4GA (Kraisig, Rosélia, Welter, Haugg, Cargnin, Roos-Frantz, Sawicki, & Frantz, 2016).

## **The Microartefacts' Distributed Architecture Model for GA**

The AHMM4GA has a defined nomenclature to facilitate its integration in any architecture model, like the ADM. The AHMM4GA is the *Entity's* holistic MM and is a set of multiple coordinated GAs that correspond to various just in time processing schemes which use the same *Project's* central pool of CSAs and CSFs. The basic MM nomenclature that the base for AHMM4GA, 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 MM's basics for AHMM4GA. The DMS4GA uses an AHMM4GA's instance to solve a *Project* problem.

Figure 7. The applied MM's basics nomenclature  
(Trad, 2020a).

Basic MM's Nomenclature		
<i>Iteration</i>	= An integer variable that denotes a <i>Project/ADM</i> iteration	
microRequirement	= KPI	(1)
CSF	= $\sum$ KPI	(2)
CSA	= $\sum$ CSF	(3)
Requirement	= $\bigcup$ microRequirement	(4)
microKnowledgeArtefact	= $\bigcup$ knowledgeItem(s)	(4)
neuron	= action ° data + microKnowledgeArtefact	(5)
microArtefact	= $\bigcup$ (e)neurons	(6)
microEntity or Enterprise	= $\bigcup$ microArtefact	(7)
Entity or Enterprise	= $\bigcup$ microEntity	(8)
microArtefactScenario	= $\bigcup$ microArtefactDecisionMaking	(9)
Decision Making/Intelligence	= $\bigcup$ microArtefactScenario	(10)
EntityIntelligence	= $\bigcup$ Decision Making/IntelligenceComponent	(11)
MM( <i>Iteration</i> ) as an instance	= EntityIntelligence( <i>Iteration</i> )	(12)

The proposed *Entity's* architecture/structure and the management of MMs enables the possibility to define the AHMM4GA; using CSFs weightings and ratings, based on multicriteria evaluation. The symbol  $\sum$  indicates summation of all the relevant named set members, while the indices and the set cardinality have been omitted. The proposed MM 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.
- The symbol  $\sum$  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.
- The symbol  $\bigcup$  indicates sets union.
- The proposed AHMM4GA enables the possibility to define *Project* AG as a model; using CSFs weightings and ratings evaluation.
- The selected corresponding weightings to: CSF  $\in \{ 1 \dots 10 \}$ ; are integer values, that are presented in tables. The rules were presented in the RDP section.
- The selected corresponding ratings to: CSF  $\in \{ 0.00\% \dots 100.00\% \}$  are floating point percentage values.

## The Applied Holistic Mathematical Model's Structure for GA

The AHMM4GA'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; in the case of AHMM4GA they are microartefacts and not tables (Lockwood, 1999).
- In the behavioural view, these actions are designed using a set of mathematics nomenclature, the implementation of the AHMM4GA 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 *Project* GA requests.

## Entity/Enterprise Architect as an Applied Mathematical Model

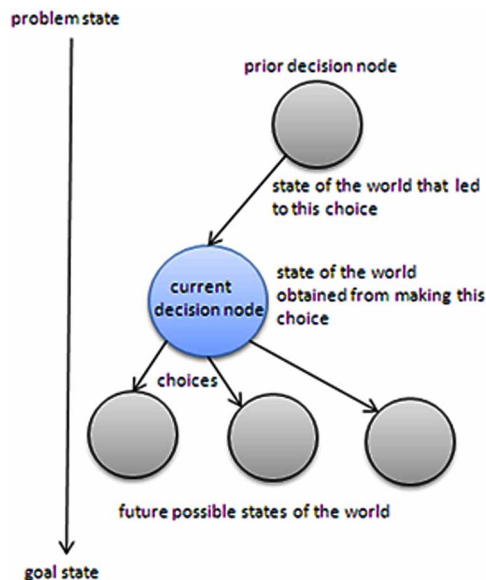
A generic *Entity* architecture model and its ADM are the kernel of this research and they are the basics of its *TRADf*. The author wants to propose a MM to represent the GA delivers an analysis. The literature review has shown that existing research resources on GA, as a MM, are practically inexistent. This pioneering research work is cross-functional and links all the GA microartefacts to *Entity* (Agievich, 2014); where the main reasoning component is a qualitative engine that is based on heuristics.

## Heuristics and Action Research

The AHMM4GA is based on a set of synchronized MM instances, where each AHMM4GA can launch a qualitative beam-search based heuristic processing (Kim & Kim, 1999; Della Croce & T'kindt, 2002). Weightings and ratings concept support the AHMM4GA to process a GA request for an optimal analysis or solution for a given *Entity* problem. Actions research can be considered as a set of continuous beam-search heuristics processing phases and is similar to design, analysis and architecture processes, like the ADM (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 GA actions useless and complex to implement. The AHMM4GA is responsible for the qualitative heuristic process for *Entity's* problem solving and synchronizes a set of MMs 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 AHMM4GA capacities are measured by analysing the *TRADf's* AHMM4GA tree.

AR based heuristics enables reflective practice that is the basis of a holistic approach to develop EA based GAs, where its kernel and skeleton are a dynamic

Figure 8. The applied heuristics tree algorithm  
(Nijboer, Morin, Carmien, Koene, Leon & Hoffman, 2009).



DMS4GA (Leitch & Day, 2006). Such an GA is based on both qualitative and quantitative methods (Loginovskiy, Dranko, & Hollay, 2018).

Qualitative, Quantitative and the Notion of Time

As already mentioned, the AHMM4GA and its underlining set of MM instances is mainly a qualitative beam-search heuristic tree (Della Croce & T’kindt, 2002). In each of the tree’s node a precise call to a quantitative functions (or other) can be executed, by precision or objectivity the author refers to input data, constraint (and/or rules) and above all a timestamp tracing system. These form the basis of an applied GA transformation based on MM instances.

Figure 9.

The Generic AHMM's Formulation

$$\text{AHMM} = \bigcup \text{ADM}s + \text{MM}s$$

(13)

The Applied GA Transformation Mathematical Model

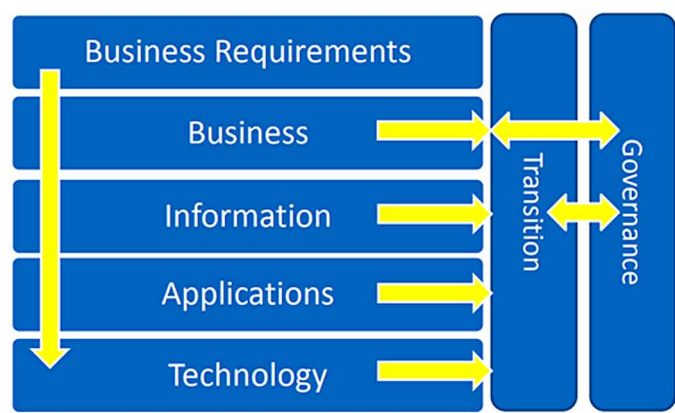
A holistic GA is a part of the *TRADf* that uses microartefacts to support just-in-time DMS4GA actions. The GA component and interface, are based on a light version of the ADM, having a system’s approach. The *Entity*’s transformation is the combination of an GA, EA methodology (like the TOGAF’s ADM) and the proposed generic AHMM, that is presented in Figures 10 and 11.

Figure 10. The MM generic structure.

AHMM's Application and Instantiation for a Specific Domain		
Domain	= Geopolitical Analysis (GA)	(14)
AHMM(Domain)	= $\bigcup$ ADMs + MMs(Domain)	(15)

The generic AHMM can applied to any specific domain; in this chapter and RDP’s phase, the *Domain*, is GA and the AHMM4GA = AHMM(GA), as shown in Figure 11.

Figure 11. The AHMM generic structure



The proposed combination can be modelled after the following formula for the Geopolitics Transformation Mathematical Model (GTMM) that abstracts the *Project* for a given *Entity*:  
(AHMM4GA for an *Iteration*):

$iAHMM = AHMM4GA(Iteration);$

$$iAHMM = Weigthing_1 * iAHMM\_Qualitative + Weigthing_2 * iAHMM\_Quantitative \quad (1)$$

$$\text{The } Project's \text{ AHMM4GA (PAHMM4G)} = \sum iAHMM \text{ for an ADM's instance} \quad (2)$$

(GTMM):

$$GTMM = \sum PAHMM4G \text{ instances} \quad (3)$$

The Main Objective Function (MOF) of the GTMM's formula can be optimized by using constraints and with extra variables that need to be tuned using the AHMM4GA. The variable for maximization or minimization can be, for example, the *Project* success, costs or other (Dantzig, 1949; Sankaralingam, Ferris, Nowatzki, Estan, Wood & Vaish, 2013). For this PoC the success will be the main and only constraint and success is quantified as a binary 0 or 1. Where the MOF definition will be:

$$\text{Minimize risk TMM} \quad (4)$$

The AHMM4GA is based on a concurrent and synchronized *TRADf*, which uses concurrent threads that can make various MMs run in parallel and manage information through the use of the AHMM4GA's mathematical choreography/ language. The GTMM is the combination of an GA, *Project* methodologies and a holistic MM that integrates the *Entity* or enterprise organisational concept, information and communication technologies that have to be formalized using a functional development environment or an NLP.

## Functional Development

*TRADf*'s internal NLP tool and its ML can be used for various application domains and in general for hard system's thinking. The author recommends the use of an interpretable scripting for building a GA (Moore, 2014; North, 2010). The AHMM4GA based process is domain-driven and is founded on *TRADf* that in turn is based on a ML to manage heuristics/rules, *Entity*/enterprise architecture and ICS microartefacts (The Open Group, 2011a; Simonin, Bertin, Traon, Jezequel & Crespi, 2010). The complexity lies in how to integrate the AHMM4GA and its programming ML in a GA and the related ICS.

*Table 2. The critical success factors that have an average of 8.90.*

As shown in Table 2, the result's aim is to prove or justify that it is complex but possible to implement a MM in the information system. The next CSA to be analysed is the holistic management of the ICS category.

## THE APPLIED CASE STUDY

## Information, Communication and Architecture Infrastructure

- Offer a GA set of possible solutions and design GA applicable patterns.
- Select a set of objectives from the proposed set of CSAs.
- Build GA microartefacts to support the various types of GA scenarios.
- Prepare *TRADf*'s phase one and if successful, select a problem from the ACS to prove phase two.



## **Integrating Critical Success Factors**

A CSF its KPI enumerations are measurable and mapped to a weighting that is roughly estimated in the first iteration and then tuned through ADM iterations, to support the GA; where a holistic set of CSFs are essential (Felfel, Ayadi, & Masmoudi, 2017); this process of evaluation is described in this RDP (Trad & Kalpić, 2019c). The main issue here is how to define the GA's goals to integrate DMS4GA and how to interrelate the different fields like culture, finance, ICS and other.

## **The Architecture Development Method and Projects**

Currently, distributed intelligence, complexity, knowledge, economy, complex business models and technology, need an ADM supported by a hyper-heuristics tree that supports a wide class of problem types that are processes in the RDP's Phase 2 (Markides, 2015), where the ADM supports GA teams. The *TRADf*'s parts must synchronize with the ADM, where the GA and its internal components are interfaced in all the ADM phases.

## **GA Aspects**

The applied GA case study analyses the fragility and recommends how to build an economically resilient Lebanon. Where economic resilience is beyond economic growth, it is about transforming the structures and face challenges. Manage banking relationships with various countries protect its financial system; that should to supported by a legal framework (OECD).

## **The Business Case Study for Architecture Critical Success Factors and the Link to the Next Applied Mathematical Model Section**

Based on the CSF review and evaluation processes, the evaluation is done with the relation and influence of architecture to this section. The important business case's CSFs that are used and processed by the internal heuristic engine; it is strongly recommended to refer to this chapter. Based on the CSF review process, the important business case's CSFs are used and evaluated as explained below.

## THE INFORMATION AND COMMUNICATION TECHNOLOGY SYSTEM

## Evolution, Development, Operations, Choreography and Maintenance

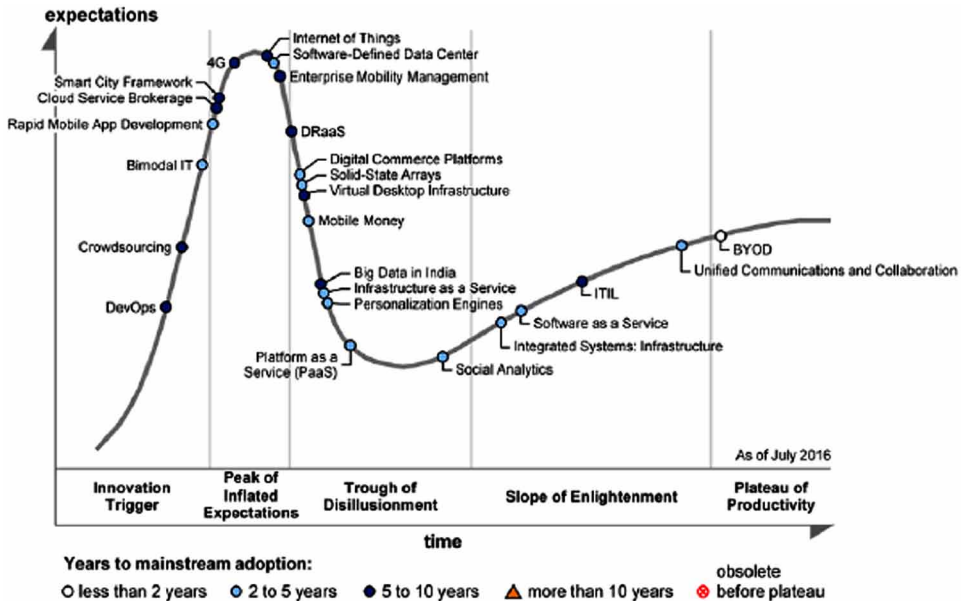
Actual methodologies/architecture, modelling, development, operations, integration and transformation tools/environments are skeletons that should enclose various automated ML microartefacts building/assembling capabilities, needed in a holistic and unified implementation strategy for a GA. Such tools must respect and adapt existing components implementation standards (Kraisig, Rosélia, Welter, Haugg, Cargnin, Roos-Frantz, Sawicki, & Frantz, 2016). The AHMM4GA formalism is based on existing proven standard methodologies, architectures and paradigms that are based on services concepts to support the GA design process, implementation and choreography.

## The Design First Approach

Defining a ML microartefact granularity and responsibility for a *Project* are a very complex undertaking in the holistic implementation of the *Project*. The design first approach supports the *TRADf's* AHMM4GA microartefacts design (Neumann, 2002). As shown in Figure 12, the AHMM4GA based process offers an graphical user interface to manage the automated and auto-generated build and deploy formalism.

The AHMM4GA formalism expresses a holistic structural concept or schema for the *Project's* GA's capabilities.

Figure 12. The graphical user interface for development and operation client interface.



## A Holistic Microartefacts Delivery Model

ML microartefacts' manipulation and its contained intelligence is in fact, a set of micro-actions that manage various GA based activities. The AHMM4GA structure is used to orchestrate ML microartefact instances and receives/evaluates change requests. The AHMM4GA includes an ML to manage interaction with other *Project's* microartefacts. The AHMM4GA's concept is based on a holistic systemic approach to use all the *TRADf's* ML microartefacts using an agile implementation process (Daellenbach & McNickle, 2005).

## A Holistic n-tier Architecture

The integration of AHMM4GA based process in the Gm, is the backbone of the future n-tiered decoupled *Entity* system (Loginovskiy, Dranko, & Hollay, 2018). An adaptable, tuneable and cross-functional AHMM4GA formalism is important for the future of any *Entity* or organisational ICS and a holistic integration strategy has to be defined using a standardized methodology like TOGAF and its Archimate modelling environment for GA (Vicente, Gama & Mira da Silva, 2013). The AHMM4GA formalism fits in the ADM (Tripathy & Mishra, 2017; Greefhorst, 2009).

## Holistic Tests, Performance, Integration and Monitoring

The major problem that causes a *Project* and GAs to be stopped or to fail, is the intelligence capacity problem that in general in *Entities* and business enterprises in general, is translated and justified by the human behavioural and cultural aspects; that is the major reason for the emergence of the saviour's new mirage, like Fowler's Microservices and astonishingly again with the same mammoth approach with the goal of grabbing more money...

Figure 13. The decision making maturity evolution (Gartner, 2016).

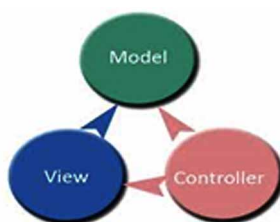


Figure 13, shows that actual immature development and operations for GAs is still in an infancy age and enterprises are losing a lot of energy on putting *Projects* and complex analysis together. These problems show that existing tools are still

*Table 4. The critical success factors that have an average of 9.20.*

		Critical Success Factors	KPIs	Weightings Ranges	Values
6	1	CSF_ICS_GUID_IntegrationProcessesModels	Standard	From 1 to 10.	10
7	2	CSF_ICS_TRADf_StandardsIntegration	AdvancedState	From 1 to 10.	10
8	3	CSF_ICS_Microartefacts_for_GA	Supported	From 1 to 10.	8
9	4	CSF_ICS_Performance_Support	Exists	From 1 to 10.	9
10	5	CSF_ICS_DistributedCommunication	Stable	From 1 to 10.	10
11	6	CSF_ICS_GA_Processing	Standard	From 1 to 10.	8
12	7	CSF_ICS_Security	Supported	From 1 to 10.	9
13	8	CSF_ICS_Automation	Supported	From 1 to 10.	9
14	9	CSF_ICS_Pattern_StandardsIntegration	Supported	From 1 to 10.	9
15	10	CSF_ICS_Procedures	Supported	From 1 to 10.	10

Eval ICS

RESULT: 9.2

< > | Entity | RDP | AMM | ACS | **ICS** | ADM | FIN | HR | KMS | DMS | GA | P1 | (+)

immature for large enterprise intelligent applications and hence *Projects* (Gartner, 2016).

## **The Mathematical Model's Integration in the Information and Communication Technology's Critical Success Factors**

Based on the literature review process, the most important ICS's CSFs that are used are evaluated to the following:

As shown in Table 4, the result tries to prove or justify that it is complex but possible to implement a MM in the information and communication system. The next CSA to be analysed is the holistic management of the GA category.

## **THE INTEGRATION WITH THE ARCHITECTURE DEVELOPMENT METHOD**

Actual archaic methodologies are inadaptable for GA, where minimum support exists for modelling the underlying goals of a GA, mainly in terms of stakeholder goals and concerns. These high-level, purely bookkeeping-financial goals, address these commercial concerns, only. This chapter's focus is on the GA concept that supports the mapping, development and modelling of linking microartefacts to all GA resources. GA's concept is based on existing standards and frameworks for goals and microartefacts mapping/modelling and is aligned with existing ADM.

## **The Unit of Work-An Extreme Granular Approach**

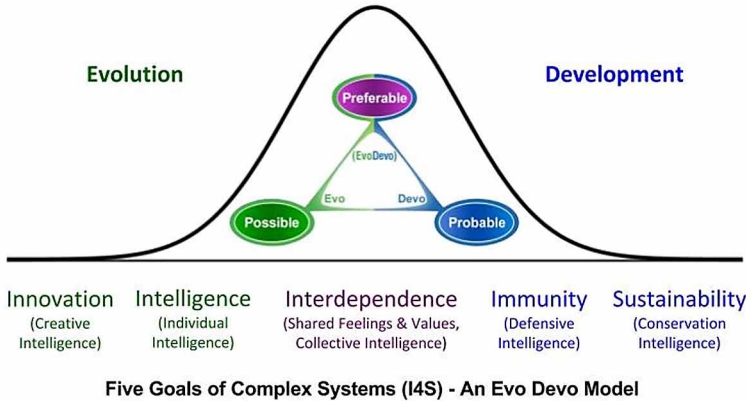
Defining an microartefact granularity and responsibility for a *Project* is a complex process; added to that, there is the complexity in implementing the "1:1" mapping and classification of the discovered microartefacts. The applied GA uses standard design methodologies like the TOGAF's ADM. This proposed GA's design and mapping concepts are supported by a set of the *TRADf*'s microartefacts where its internal AHMM4GA NLP consists of implementing microartefacts to dynamically evaluate compound expressions, according to the AHMM4GA principles (Neumann, 2002).

## **The Model First Approach-Top Down Approach**

The pseudo bottom up approach of an strategy used by the *TRADf* is influenced by the microartefacts that are managed by the GA strategy, methodology and productivity environment. The *TRADf* proposes an upstream concept that are altered to accommodate traditional services environments; these services are stored into

a specialized atomic service repository using the model-view-control pattern, as shown in Figure 14. The author recommends that *Managers* should apply a EAP as a base for their *Project* strategy.

*Figure 14. The MVC.*



## Holistic Integration Architecture

Architectures derived from standardized enterprise architectures, like the TOGAF differ greatly, because they depend on GA's requirements quality. In implementing services based architecture *TRADf* can be used to support and assess the requirements. The *TRADf* focus is on the development and modelling of a holistic architectures.

## The Architecture Development Method Critical Success Factors and Link to the Next Section

Based on the literature review and associated evaluation processes, the CSFs are evaluated (Trad & Kalpić, 2019e) and are explained below.

## THE HOLISTIC STRATEGIC HUMAN RESOURCES SYSTEM

In the previous phases of the RDP has concluded that a *Manager* is Architect of Adaptive Business Information Systems (AofABS), specialized in managing the implementation phase of a *Project* (Trad, & Kalpić, 2020a).

This RDP presents an original set of CSFs and fulfils the need for an efficient GA based system. The sets of CSFs are presented in the form of a real-world constraints, which affect the *Manager* selection techniques. *Manager* for GA selectors, professional analysts, project managers, auditors and advanced computer science students might benefit from this research project, while its ambition is to be considered as a major managerial benefit.

Actually, there is no concrete educational curriculum for a *Manager* for GA, where mainly bookkeeping profiles are chosen for such positions. There is an essential need for more research on the *Managers'* profiles and their educational prerequisites (Tidd, 2006). The needed skills must comprise the knowledge of: 1) business and enterprise architecture; 2) automated real-time business process environments; 3) agile project management; 4) organizational behaviour; 5) management sciences

methodologies; and 6) concrete ICS implementation phase know-how and experience. As already mentioned, the researcher recommends, the profile of an AofABS (Trad, & Kalpić, 2020a).

## TRADf's Support

The GA oriented profile's selection can be supported by *TRADf* which uses a hyper-heuristics reasoning model for the selection process. The reasoning model, is a qualitative reasoning tree to support to select the *Manager's* profile, with the defined constraints These revealed CSA/CSFs are also fundamental for the future coordination of *Projects*. The DMS4GA will offer the relationships between the different CSAs in order to rate and weight and rate the CSFs.

## The Technocrat's Profile

*Managers* who are basically (geopolitical) technocrats who should be capable of supporting and designing a GA module (Farhoomand, 2004); they need cross-functional skill. The profile's selected CSFs are fundamental for the selection of the *Manager*, where it develops a concept for the profile's selection.

## A Holistic Approach

*Managers* need more than basic DMS4GA, KMS4GA and ICS support to exploit GA in order to successfully conduct the *Project*. Such *Managers* and organizations

Table 6. The critical success factors that have an average of 9.80.

	Critical Success Factors	KPI	Weightings Ranges	Values
1	CSF_FIN_BanksInfluence	Confirmed	From 1 to 10.	10
2	CSF_FIN_Brutal_RuthlessActs	Highly Probable	From 1 to 10.	9
3	CSF_FIN_Misdeeds	Confirmed	From 1 to 10.	10
4	CSF_FIN_GeopoliticalInfluence	Confirmed	From 1 to 10.	10
5	CSF_FIN_FinancialCrimes	Confirmed	From 1 to 10.	10

Eval FIN

RESULT: 9.8



need a framework, like *TRADf*, which encompasses various fields. This RDP shows that the *Manager* is an AofABS with cross-functional skills (Uhl, & Gollenia, 2012).

## Educational Requirements

The *Managers* need hands-on skills and educational requirements that encompasses the following set of skills: 1) GA related knowledge of business architectures and business processes; 2) automated complex analysis environments; 3) agile project management; 4) knowledge management & integration; 5) *Entity* and organizational concepts; 6) management sciences methodologies; 7) enterprise architecture and other concrete *Project* implementation artefacts.

## The Human Resources Success Factors

Based on the literature review and associated evaluation processes, the CSFs are evaluated and are explained below.

*Table 7. The critical success factors that have an average of 9.20.*

		Critical Success Factors	KPIs	Weightings Ranges	Values
1		CSF_HR_RDP	Standardized	From 1 to 10.	10
2		CSF_HR_GSA_CSF	AdvancedState	From 1 to 10.	10
3		CSF_HR_Surveying	Supported	From 1 to 10.	8
4		CSF_HR_SkillsProfile	Exists	From 1 to 10.	9
5		CSF_HR_TRADf	Stable	From 1 to 10.	9
6		CSF_HR_GA_TechnocratProfile	Standard	From 1 to 10.	7
7		CSF_HR_HolisticApproach	Supported	From 1 to 10.	10
8		CSF_HR_GA_EducationalRequirements	Supported	From 1 to 10.	10
9		CSF_HR_GA_Recommendations	Supported	From 1 to 10.	10

Exec HR

RESULT: 9.2222222

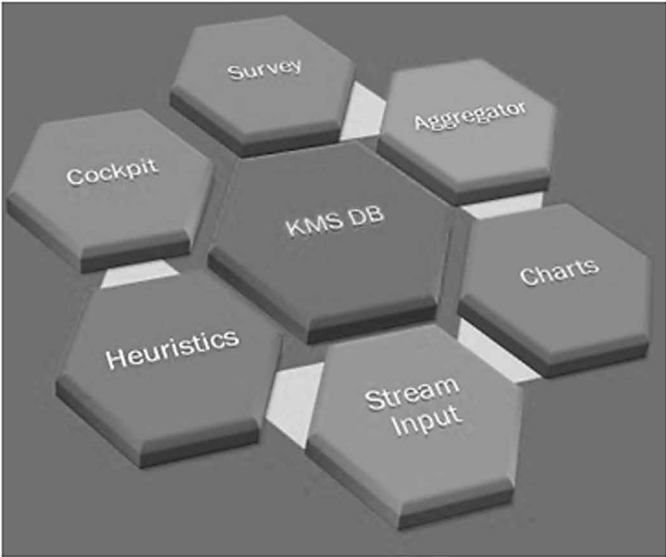
Entity RDP AMM ACS ICS ADM FIN **HR** KMS DMS GA P1

## HOLISTIC DECISION AND KNOWLEDGE MAKING SYSTEMS

### Complex Systems

Complex systems management can be adapted to the *Project*'s problems and requests by using AHMM4GA based DMS4GA/KMS4GA (Daellenbach & McNickle, 2005). The *Project* GA requests are processed by using the *TRADf*'s AHMM4GA instance,

*Figure 15. Complex system’s nature and approach.*



as shown in Figure 15 that in turn are based on the selected critical success areas and factors that can be used as a DMS4GA/KMS4GA which has a very complex system evolution nature.

*Figure 16. The knowledge management system.*





microartefacts are orchestrated by the AHMM4GA choreography engine. The AHMM4GA based DMS4GA' actions map to the various GA's mechanisms to deliver concrete actions. The AHMM4GA formalism is implemented in all of the *Project's* processes and the implementation of microartefacts to deliver a DMS4GA; such a set of actions can be modelled and managed by the AHMM4GA that is implemented with an experiment or a PoC (The Open Group, 2011a).

## **The Decision Making System's Critical Success Factors**

Based on the literature review process, the most important DMS4GA/KMS4GA CSFs that are used are evaluated to the following:

As shown in Tables 8 and 9, the result tries to prove or justify that it is complex but possible to implement a DMS4GA/DMS4GA to support the GA.

## **THE GA DOMAIN**

GA focuses on the geopolitical influence or even chaos that is a result of financial problems, crimes and even irregularities.

## **Finance Oriented Geopolitics**

This approach causes damages to various financial ecosystems; GA tries to predict such issues using the following types of information sources: 1) state financial misdeeds references; 1) banks' influence and related legal processes; 2) business/ financial valuation; 3) currency value manipulation; 4) global domestic growth indices; 5) financial institutions' strategic views; and 6) global financial and influence networks. These sources are needed to construct GA's sets of CSFs. Where financial misdeeds, like fraud and money laundering that damage many *Entities* (and even powerful countries, like the USA or France) can be detected by tracing the origins of the used dirty capital (Stupples, Sazonov, & Woolley, 2019).

## **Geopolitical Networked Influence**

GA must be aware of the influence of geopolitical CSFs that can be categorized as follows:

- Many CSFs, can influence geopolitical CSFs that can include: 1) the role of the local financial law enforcement agencies; like safe heavens, which totally support their banks; 2) geopolitical relations; 3) political, ethnical and cultural

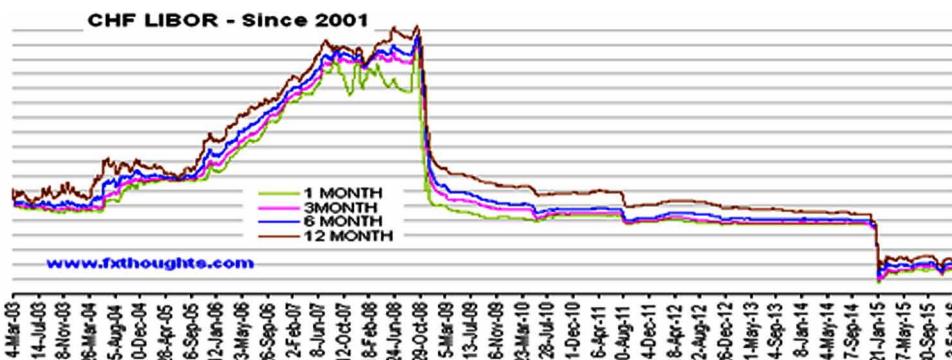
setup; 4) level of education; 5) standard of life; 6) financial competition, which in this chapter is the most important factor; and 7) financial and legal control mechanisms.

- Localise possible geopolitical frontends for financial influence that is based on elite networks.
- Lobbies builds elite networks, through good payed jobs, elite schools and financial business summits... These networks are often (mis)used for achieving financial goals.

*Figure 17. Switzerland's economy growth  
(Trading Economics, 2017a).*



*Figure 18. Lebanon's economy growth  
(Trading Economics, 2017b).*



- The influence of geopolitics can be used to destabilize fragile countries like in the case Lebanon in which its democracy is unstable; that makes it an easy target for financial predators.

Economic Growth of some countries shows the financial influence of predatory safe heavens. This section inspects if a safe heaven took advantage of the Lebanese civil war, by comparing GDP diagrams. As shown in Figures 17 and 18, analysing the GDP diagrams' slopes and it seems obvious that the Lebanese declining GDP slope is inversely equivalent to the financial safe heaven that had the aim to take takeover (Trad, 2018a).

In 1975, the Lebanon had its GDP drop to one third of its GDP of the early '70s, although its currency before the war evolved to become a worldwide solid currency. Lebanon's fast enrichment placed it in the shooting range of the financial hunters, who provoked rumour-based collapse of the Intra Bank and other Lebanese financial institutions and caused the migration of financial assets from the Lebanon to its rival banks (AMInfo, 2014). In the case of Lebanon, it is clear that the UBS mechanism is destructive.

## **State Financial Strategy**

The State Financial Strategy (SFS) is based on the following facts and assumptions:

- The notion of states applying state terrorism exists (Agger, & Jensen, 1996), so various types of means to support state terrorism exist, like religion, ideology... and in this case finance. So the term of SFS, can be labelled as states terrorising other states or/and peoples by the means of financial sabotage.
- SFS caused a fraud case worth \$2 billion USD in loans to Mozambique (Reuters, 2019a).
- In the case of Greece more 200 billion were plundered, while Greece is suffering (Le Point, 2011).
- The Nobel prize winner, the British economist, Angus Deatoun, warns about the destructive predator's graduating business schools (Le Monde, 2019). Such profiles can be classified as SFS profiles.

There are too many cases of ruthlessness of SFS that can be not all listed.

## **Global Fraud**

This section analyzes the notions of the SFS's global fraud:

- Hidden capital is reused as a credit to the already weakened countries (Stupples, Sazonov, & Woolley, 2019).
- SFS related accountancy is based on a complex locked-in legal, financial and accounting system that blurs financial flows and disables any type of transparency.
- However, some credible sources like the Global Forum on Transparency and Exchange of Information for Tax Purposes peer review in 2011, has identified important deficiencies in the legal foundations for transparency and corruption (OECD, 2011).
- In the USA, a federal judge accused the a global bank for causing *catastrophic* investor losses in residential mortgage-backed securities sold before the 2008 financial crisis that caused more than \$41 billion of damage of subprime and other risky loans in 40 offerings (Stempel, 2019).
- The financial crisis of 2007 to 2009 was marked by widespread fraud in the mortgage securitization industry (ASA, 2019).
- Paula Ramada estimated the amount of lost money due to the benchmark of interest rates debacle is estimated at \$300 trillion in financial instruments, ranging from mortgages to student loans (Trad, 2018a).
- Switzerland and its most famous banks orchestrated the deportation and dilapidation of the victims of the Second World War (Rickman, 1999).

## GA's Critical Success Factors

The most important GA CSFs that are used are evaluated to the following:

*Table 10. The critical success factors that have an average of 9.10.*

	Critical Success Factors	KPIs	Weightings Ranges	Values
6	CSF_GA_GUID_IntegrationModels	Standard	From 1 to 10.	10
7	CSF_GA_TRADf_StandardsIntegration	AdvancedState	From 1 to 10.	9
8	CSF_GA_Microartefacts_for_Scenario	Supported	From 1 to 10.	8
9	CSF_GA_RoleOfFinance	Exists	From 1 to 10.	10
10	CSF_GA_DefenseCapabilities	Stable	From 1 to 10.	10
11	CSF_GA_GA_Processing	Standard	From 1 to 10.	10
12	CSF_GA_Security	Supported	From 1 to 10.	8
13	CSF_GA_Audit	Supported	From 1 to 10.	9
14	CSF_GA_Pattern_StandardsIntegration	Supported	From 1 to 10.	9
15	CSF_GA_GovernanceProcedures	Supported	From 1 to 10.	8

Eval GA

RESULT:

9.1

## The ACS

As mentioned, the ACS is related to two cases, include an insurance claims system (ACS\_1) and a GA related case (ACS\_2).

## The PoC

The AHMM4GA related PoC was implemented using *TRADf* that had been developed by the author. The PoC uses an internal set of CSFs' that are presented in Tables 1 to 10. These CSFs have bindings to specific RDP resources, where the AHMM4GA formalism was designed using an ML microartefacts. In this chapter's tables and the result of the processing of the DMS4GA, as illustrated in Table 4, shows clearly that the AHMM4GA is not an independent component and in fact it is strongly bonded to the *Project*'s overall risk architecture, hence has to have a holistic approach.

*Table 11. The P1 outcome is 9.30.*

	Critical Success Factors	KPIs	Weightings/Ranges	Values
1	RDP	RoleOfResearch	From 1 to 10.	9,5
2	AMM	RoleOfAHMM	From 1 to 10.	8,887142857
3	ACS	AppliedCases	From 1 to 10.	8,8
4	ICS	RoleOfInfrastructure	From 1 to 10.	9,2
5	ADM	RoleOfArchitecture	From 1 to 10.	9,8
6	FIN	RoleOfFinance	From 1 to 10.	9,8
7	HR	RoleOfHumanFactor	From 1 to 10.	9,222222222
8	KMS	RoleOfKnowledgeManagement	From 1 to 10.	9
9	DMS	RoleOfDecisionMaking	From 1 to 10.	9,6
10	GA	GACapabilities	From 1 to 10.	9,1

EvaPA

RESULT:

9,2879365

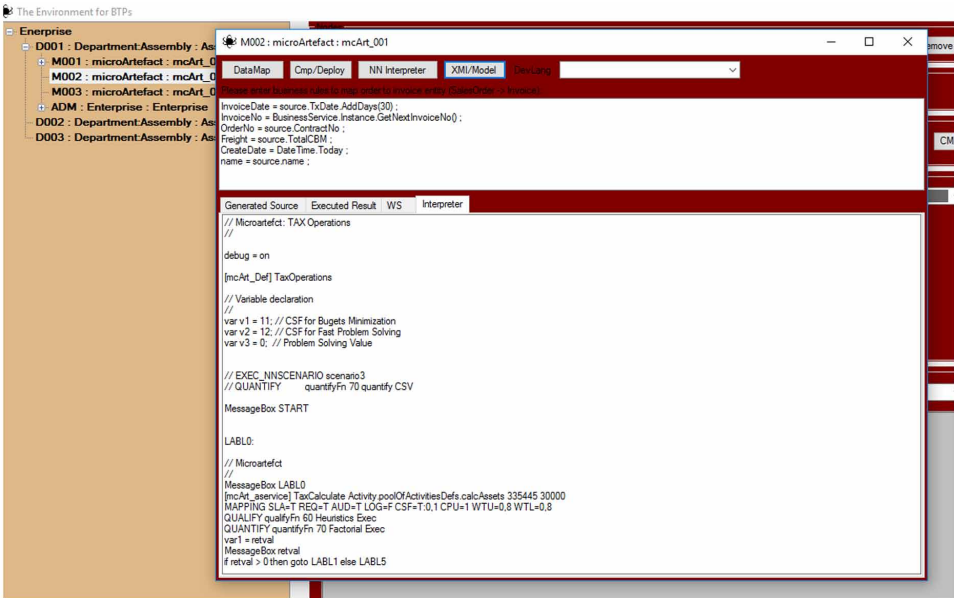
Entity RDP AMM ACS ICS ADM FIN HR KMS DMS GA P1 +

The *TRADf* and hence the AHMM4GA’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 AHMM4GA’s holistic implementation an average result below 6.5 will be ignored. As shown in Table 11, this fact keeps the CSAs (marked in green) that helps to make this work’s conclusion.

The AHMM4GA processing model represents the relationships between this research's requirements, project ML generic and microartefacts, unique identifiers

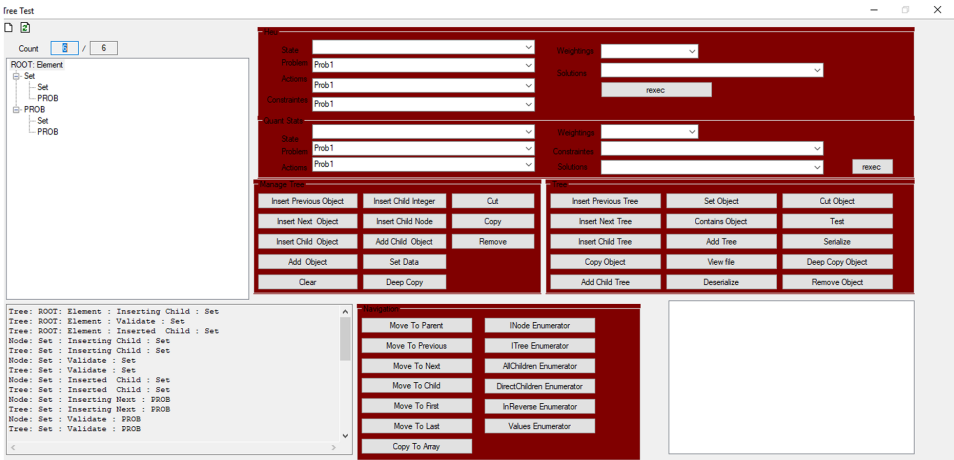


Figure 19. The edited ML script and flow



and the CSAs. The PoC was achieved using *TRADf*client's interface. From the *TRADf* client's interface the ML development setup and editing interface can be launched. Once the development setup interface is activated the NLP 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 AHMM4GA set of

Figure 20. The heuristics tree configuration.



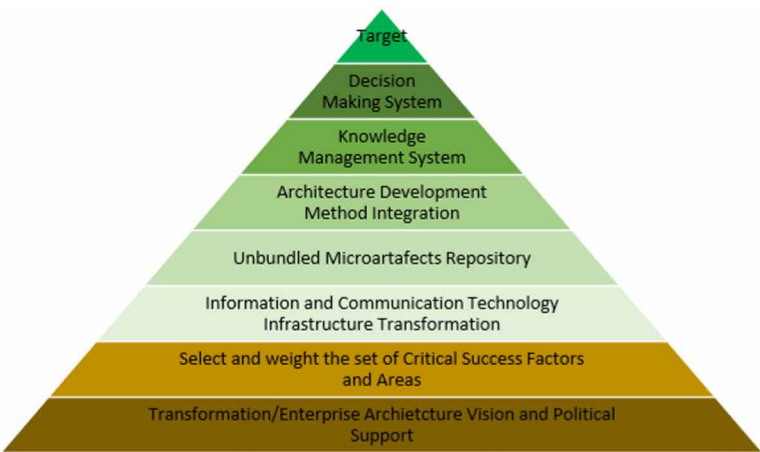
actions that are processed in the background. The AHMM4GA uses a knowledge database that automatically generates decision making actions which make calls to DMS4GA, that manages the edited ML script and flow, as shown in Figure 19.

This research’s instance of the AHMM4GA and its related CSFs were selected as demonstrated previously, as shown in Figure 20. This tree is applied on a specific CSF that is mapped to a specific problem and a set of actions.

**SOLUTION AND RECOMMENDATIONS**

Table 11 shows that AHMM4GA implementation is not a risky process; this chapter’s PoC has achieved the defined objectives. As shown in Figure 21, the AHMM4GA instance is in all of the *Project’s* processes; such a set of CSF mapped actions; like the ones presented in this chapter’s PoC.

*Figure 21. The proposed methodology and framework pyramid.*



If the aggregations of all the *Project’s* CSA/CSF tables exceeds the defined minimum, the *Project* continues to its PoC or can be used for problem solving using the heuristic algorithm with punctual calls to quantitative methods.

## FUTURE RESEARCH DIRECTIONS

The *TRADf* future research efforts will focus on the holistic integration of the chapter's mentioned various fields to increase success of transformational initiatives in a cross-functional environment.

## CONCLUSION

This RDP phase is part of a series of publications related to *Projects* which use mixed action research model; where CSFs are offered to help GA *specialists* to inspect complex situations. In this chapter, the focus is on the AHMM4GA's formalism that defines a structured inter-relationship of various MMs. To avoid major geopolitical pitfalls, the author recommends, to:

- perform *Project* operations through multiple independent sub-projects, where the priority is to transform their DMS4GA.
- A distributed DMS4GA must be built to support *Managers* on various levels.
- Existing methodologies improve the robustness of an *Entity's* infrastructure, which can unbundle ancient infrastructure.
- An *Entity* has to build an ICS based on the AHMM4GA-like concept, to counter financial misdeeds. The strategy must be built on the detection of UBS' destructive strategies.

## ACKNOWLEDGMENT

In a work as large as this research project, technical, typographical, grammatical, or other kinds of errors are bound to be present. Ultimately, all mistakes are the author's responsibility.

## REFERENCES

- Agger, I., & Jensen, S. (1996). *Trauma and Healing Under State Terrorism*. ZEB Books.
- Agievich, V. (2014). *Mathematical model and multi-criteria analysis of designing large-scale enterprise roadmap*. PhD thesis.

Alderman, L. (2019). French Court Fines UBS \$4.2 Billion for Helping Clients Evade Taxes. *The New York Times*. Retrieved October 2019, from <https://www.nytimes.com/2019/02/20/business/ubs-france-tax-evasion.html>

AMInfo. (2014). *Middle Eastern clients in the HSBC Switzerland leaks*. Swiss Leaks. Retrieved June 2017, from <http://ameinfo.com/luxury-lifestyle/list-middle-eastern-clients-in-the-hsbc-switzerland-leaks/>

ASA. (2019). *The causes of fraud in the financial crisis of 2007 to 2009: evidence from the mortgage-backed securities industry*. ASA. Retrieved October 2019, from <https://www.asanet.org/causes-fraud-financial-crisis-2007-2009-evidence-mortgage-backed-securities-industry>

Ball, R. (1968). *The Project Gutenberg eBook of a Short Account of the History of Mathematics*. Gutenberg. Dover publications.

Ball Hinkelmann, K. (2016). *Modelling in Enterprise Architecture*. University of applied sciences northwest Switzerland. Business School.

Berger, J., & Rose, J. (2015). Nine Challenges for e-Government Action Researchers. *International Journal of Electronic Government Research*, 11(3), 57–75. doi:10.4018/IJEGR.2015070104

Brown, J. (2016). *Wafic Said: businessman, philanthropist and political fixer*. Financial Times. Retrieved May 15, 2017, from <https://www.ft.com/content/a3cb764a-ecf1-11e5-bb79-2303682345c8>

Busenitz, L. (2014). *Entrepreneurial Risk and Strategic Decision Making, It's a Matter of Perspective*. SAGE Journals.

Cearley, D., Walker, M., & Burke, B. (2016). *Top 10 Strategic Technology Trends for 2017*. Academic Press.

Chandrasekhar, A. (2018). *Trial of LTTE Financiers Begins in Switzerland-The 13 on trial face charges of fraud, false documentation, money laundering and extortion*. WIRE.

Chu, F., & Xie, X. (1997). *Deadlock analysis of Petri nets using siphons and mathematical programming*. *IEEE Transactions on Robotics and Automation*, 13(6), 793 - 804.

Cornevin, C. (2020). *La police démantèle un vaste système de blanchiment de fraude fiscale... Le Figaro*. Retrieved January 2020, from <https://www.lefigaro.fr/actualite-france/la-police-demantele-un-vaste-systeme-de-blanchiment-de-fraude-fiscale-20200110>

Daellenbach, H., McNickle, D., & Dye, Sh. (2012). *Management Science - Decision making through systems thinking* (2nd ed.). Palgrave Macmillian.

Della Croce, F., & T'kindt, V. (2002). A Recovering Beam Search algorithm for the one-machine dynamic total completion time scheduling problem. *The Journal of the Operational Research Society*, 53(11), 1275–1280. doi:10.1057/palgrave.jors.2601389

Dogan, S., Çalgici, P., Arditi, D., & Gunaydin, H. (2015). *Critical success factors of partnering in the building design process. METU JFA 2015/2*. Department of Architecture, İzmir Institute of Technology.

Duparc, A. (2010). *La Suisse restitue au Liban les archives du fonds Dunand*. Le Monde.

Easterbrook, S., Singer, J., Storey, M., & Damian, D. (2008). *Guide to Advanced Empirical Software Engineering-Selecting Empirical Methods for Software Engineering Research*. Springer.

El Hashem, B. (1990). *It was Kissinger who destroyed the nation of Lebanon*. EIR Feature.

European commission, (2004). Commission of the European Communities Legal barriers in e-business: The results of an open consultation of enterprises. Brussels, 26.4.2004 SEC(2004) 498. European commission.

Felfel, H., Ayadi, O., & Masmoudi, F. (2017). Pareto Optimal Solution Selection for a Multi-Site Supply Chain Planning Problem Using the VIKOR and TOPSIS Methods. *International Journal of Service Science, Management, Engineering, and Technology*. Doi:10.4018/IJSSMET.2017070102

Fitsanakis, J. (2016). *Switzerland made secret deal with PLO in the 1970s, new book alleges*. Academic Press.

Gartner. (2013). *Scenario Toolkit: Using EA to Support Business Transformation*. Gartner Inc.

Gartner. (2016). *Gartner's 2016 Hype Cycle for ICT in India Reveals the Technologies that are Most Relevant to Digital Business in India Analysts to Explore Key Technologies and Trends*. Gartner Symposium/ITxpo 2016, Goa, India.

Giachetti, R. (2012). *A Flexible Approach to Realize an Enterprise Architecture*. Department of Systems Engineering, Naval Postgraduate School, Monterey, CA USA. Presented: New Challenges in Systems Engineering and Architecting. Conference on Systems Engineering Research (CSER). St. Louis, MO. 10.1016/j.procs.2012.01.031

- Goikoetxea, A. (2004). A mathematical framework for enterprise architecture representation and design. *International Journal of Information Technology & Decision Making*, 03(01), 5–32. doi:10.1142/S0219622004000623
- Greefhorst, D. (2009). *Using the Open Group's Architecture Framework as a pragmatic approach to architecture*. Jaarbeurs, Utrecht. KIVINIRIA, afd. Informatica.
- Gunasekare, U. (2015). *Mixed Research Method as the Third Research Paradigm: A Literature Review*. University of Kelaniya.
- Ho, W., Xu, X., & Dey, P. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. Operations and Information Management Group, Aston Business School, Aston University.
- Izzo, S. (2019). Karl-Heinz Hoffmann's Secret History Links Neo-Nazis With Palestinian Terror - Tablet Magazine. *Tablet (Brooklyn, N.Y.)*. Retrieved October 2019, from <https://www.tabletmag.com/jewish-arts-and-culture/culture-news/286220/karl-heinz-hoffmann-far-right> 1/11
- Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality & Quantity*, 41(1), 37–54. Retrieved August 10, 2018, from <https://link.springer.com/article/10.1007/s11135-005-5427-1>
- Johnson, R., & Onwuegbuzie, A. (2004). *Mixed Methods Research: A Research Paradigm Whose Time Has Come*. Sage Journals.
- Johnson, S. (1983). Francois Genoud: Terrorist controller for Swiss banks. *Executive Intelligence Review*.
- Jonkers, H., Band, I., & Quartel, D. (2012a). *ArchiSurance Case Study*. The Open Group.
- Joseph, C. (2014). *Types of eCommerce Business Models*. Retrieved September 17, 2019, from <https://smallbusiness.chron.com/types-ecommerce-business-models-2447.html>
- 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
- Kraisig, A., Rosélia, A., Welter, F., Haugg, I., Cargnin, R., Roos-Frantz, F., Sawicki, S., & Frantz, R. (2016). Mathematical Model for Simulating an Application Integration Solution in the Academic Context of Unijuí University. *Procedia Computer Science*, 100, 407–413. doi:10.1016/j.procs.2016.09.176

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

Le Monde. (2019). *Le Prix Nobel d'économie Angus Deaton: « Quand l'Etat produit une élite prédatrice »*. Le Monde. Retrieved November 14, 2019, from [https://www.lemonde.fr/idees/article/2019/12/27/angus-deaton-quand-l-etat-produit-une-elite-predatrice\\_6024205\\_3232.html](https://www.lemonde.fr/idees/article/2019/12/27/angus-deaton-quand-l-etat-produit-une-elite-predatrice_6024205_3232.html)

Le News. (2015). *Swiss People's Party (UDC) leaders found guilty of racism*. Le News. Retrieved September 2019, from <https://lenews.ch/2015/04/30/two-swiss-peoples-party-udc-leaders-found-guilty-of-racism/>

Le News. (2017). *Racism sentence upheld against former Swiss People's Party secretary general*. Le News. Retrieved September 2019, from <https://lenews.ch/2017/04/13/racism-sentence-upheld-against-former-swiss-peoples-party-secretary-general/>

Leitch, R. & Day, C. (2000). *Action research and reflective practice: towards a holistic view*. Taylor & Francis.

Lockwood, R. (2018). *Introduction The Relational Data Model*. Retrieved January 29, 2019, from <http://www.jakobsens.dk/Nekrologer.htm>

Loginovskiy, O. V., Dranko, O. I., & Holloy, A. V. (2018). *Mathematical Models for Decision-Making on Strategic Management of Industrial Enterprise in Conditions of Instability*. Conference: Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications (EMIT 2018). Leipzig, Germany.

Markides, C. C. (2015). Research on Business Models: Challenges and Opportunities. *Advances in Strategic Management*, 33, 133–147. doi:10.1108/S0742-332220150000033004

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

Moore, J. (2014). *Java programming with lambda expressions-A mathematical example demonstrates the power of lambdas in Java 8*. Retrieved March 10, 2018, from, <https://www.javaworld.com/article/2092260/java-se/java-programming-with-lambda-expressions.html>

- Myers, B., Pane, J., & Ko, A. (2004). *Natural programming languages and environments*. ACM New York. doi:10.1145/1015864.1015888
- Neumann, G. (2002). *Programming Languages in Artificial Intelligence*. In *Encyclopaedia of Information Systems*. Academic Press.
- 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
- Nilda Tri, P., & Yusof, S. M. (2009). Critical Success Factors for Implementing Quality Engineering Tools and Techniques in Malaysian's and Indonesian's Automotive Industries: An Exploratory Study. In *Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2009*. MECS.
- North, N. (2010). *Behaviour-Driven Development Writing software that matters*. DRW publications.
- OECD. (2011). *Global Forum on Transparency and Exchange of Information for Tax Purposes Peer Review: Switzerland 2011, Phase 1*. OECD Publishing.
- OECD. (2018). *Country Case Study 1: Lebanon. MENA-OECD Economic Resilience Task Force Resilience In Fragile Situations. 4-5 December 2018. Islamic Development Bank*. OECD.
- Paravicini, G. (2018). *Millions flow from Gaddafi's 'frozen funds' to unknown beneficiaries*. Politico. <https://www.politico.eu/article/muammar-gaddafi-frozen-funds-belgium-unknown-beneficiaries/>
- Peterson, S. (2011). *Why it Worked: Critical Success Factors of a Financial Reform Project in Africa*. Faculty Research Working Paper Series. Harvard Kennedy School.
- Polderman, J., & Willems, J. (1998). *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer Verlag, Germany. doi:10.1007/978-1-4757-2953-5
- Ravanetti, A. (2016). *Switzerland Bank on Fintech with Lighter Regulations*. Crowd Valey. Retrieved September 2019, from <https://news.crowdvalley.com/news/switzerland-bank-on-fintech-with-lighter-regulations>
- Reuters. (2019a). *Swiss group files criminal complaint against Credit Suisse over Mozambique loans*. Reuters. <https://www.reuters.com/article/us-mozambique-creditsuisse/swiss-group-files-criminal-complaint-against-credit-suisse-over-mozambique-loans-idUSKCN1S5174>



- Rickman, R. (1999). *Swiss Banks and Jewish Souls* by Gregg J. Rickman. Central European History. JSTOR.
- Sankaralingam, K., Ferris, M., Nowatzki, T., Estan, C., Wood, D., & Vaish, N. (2013). *Optimization and Mathematical Modeling in Computer Architecture*. Morgan & Claypool Publishers.
- Scherer, R. J., & Schapke, S. E. (2011, October). A distributed multi-model-based Management Information System for simulation and decision making on construction projects. *Advanced Engineering Informatics*, 25(4), 582–599. doi:10.1016/j.aei.2011.08.007
- 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
- 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
- Snowden, E. (2015). *Most Racist, Award Goes To ... Switzerland? Skating on Stilts*. Retrieved September 2019, from <https://www.skatingonstilts.com/skating-on-stilts/2015/03/and-the-edward-snowden-most-racist-award-goes-to-switzerland.html>
- Stempel, J. (2019). *UBS must defend against U.S. lawsuit over 'catastrophic' mortgage losses*. Yahoo Finance. Retrieved September 2019, from <https://finance.yahoo.com/news/ubs-must-defend-against-u-214743943.html>
- Stupples, B., Sazonov, A., & Woolley, S. (2019). *UBS Whistle-Blower Hunts Trillions Hidden in Treasure Isles*. *Bloomberg-Economics*. Bloomberg. Reviewed in November 2019 <https://www.bloomberg.com/news/articles/2019-07-26/ubs-whistle-blower-hunts-trillions-hidden-in-treasure-islands>
- Syynimaa, N. (2015). *Enterprise Architecture Adoption Method for Higher Education Institutions* (Doctoral Thesis). Informatics Research Centre Henley Business School University of Reading.
- The Open Group. (2011a). *Architecture Development Method*. The Open Group. USA. Reviewed in February 2018, <https://pubs.opengroup.org/architecture/togaf9-doc/arch/chap05.html>

- The Open Group. (2011a). *Architecture Development Method*. The Open Group. Reviewed in February 2018, <https://pubs.opengroup.org/architecture/togaf9-doc/arch/chap05.html>
- The Open Group. (2011b). *TOGAF 9.1*. The Open Group. Reviewed in August 2018, <http://www.opengroup.org/subjectareas/enterprise/togaf>
- Thomas, A. (2015). *Gartner, Innovation Insight for Microservices*. Gartner.
- Thomas, A. (2015). *Gartner, Innovation Insight for Microservices*. Gartner.
- Tidd, J. (2006). *From Knowledge Management to Strategic Competence* (2nd ed.). Imperial College. doi:10.1142/p439
- Tidd, J., & Bessant, J. (2009). *Managing Innovation, Integrating Technological, Market and Organizational Change* (4th ed.). Wiley.
- Trad, A. (2018a). *The Business Transformation and Enterprise Architecture Framework-Applied to analyse / The historically recent Rise and the 1975 Fall of the Lebanese Business Ecosystem*. IGI-Global.
- Trad, A. (2019d). *Applied Mathematical Model for Business Transformation-Assessing Risks of the Lebanese Islamic Business/Marketing Strategy and its Relationships with Counterparts/Partners (IBM&R)*. IGI-Global.
- Trad, A. (2019e). *Applied Mathematical Model for Business Transformation Projects-The intelligent Strategic Decision Making System (iSDMS)*. Encyclopaedia. IGI-Global.
- Trad, A., & Kalpić, D. (2017b). *An Intelligent Neural Networks Micro Artefact Patterns' Based Enterprise Architecture Model*. IGI-Global.
- Trad, A., & Kalpić, D. (2017c). *A Neural Networks Portable and Agnostic Implementation TKM&F for Business Transformation Projects. The Framework*. IEEE.
- Trad, A., & Kalpić, D. (2017d). *A Neural Networks Portable and Agnostic Implementation TKM&F for Business Transformation Projects- The Basic Structure. IEEE Conference on Computational Intelligence*.
- Trad, A., & Kalpić, D. (2018a). *The Business Transformation Framework and Enterprise Architecture Framework for Managers in Business Innovation-Knowledge and Intelligence Driven Development (KIDD). Encyclopaedia of E-Commerce Development, Implementation, and Management*. IGI-Global.

Trad, A., & Kalpić, D. (2018b). *The Business Transformation Framework and Enterprise Architecture Framework for Managers in Business Innovation- Knowledge Management in Global Software Engineering (KMGSE)*. *Encyclopaedia of E-Commerce Development, Implementation, and Management*. IGI-Global.

Trad, A., & Kalpić, D. (2019c). *Business Transformation and Enterprise Architecture- The Resources Management Research and Development Project (RMSRDP)*. Book. IGI-Global.

Trad, A., & Kalpić, D. (2019e). *Business Transformation and Enterprise Architecture- The Holistic Project Resources Management Pattern (HPRMP)*. *Encyclopaedia*. IGI-Global.

Trad, A., & Kalpić, D. (2020a). *Using Applied Mathematical Models for Business Transformation*. *IGI Complete Author Book*. IGI Global. doi:10.4018/978-1-7998-1009-4

Trading Economics. (2017a). *Switzerland - GDP Annual Growth Rate*. Trading Economics. April 10, 2017, from <http://www.tradingeconomics.com/switzerland/gdp-growth-annual>

Trading Economics. (2017b). *Lebanon - GDP Annual Growth Rate*. Trading Economics. Retrieved April 10, 2017, from <http://www.tradingeconomics.com/lebanon/gdp-growth-annual>

Tripathy, B., & Mishra, J. (2017). *A Generalized Framework for E-Contract*. *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*. IGI Global., doi:10.4018/IJSSMET.2017100101

Türkmen, E., & Soyer, A. (2020). The Effects of Digital Transformation on Organizations. In *Handbook of Research on Strategic Fit and Design in Business Ecosystems* (pp. 259-288). IGI Global. doi:10.4018/978-1-7998-1125-1.ch011

Uhl, L., & Gollenia, L. A. (2012). *A Handbook of Business Transformation Management Methodology*, Gower. SAP.

Zaiane, S., & Ben Moussa, F. (2018). *Cognitive Biases, Risk Perception, and Individual's Decision to Start a New Venture*. *International Journal of Service Science, Management, Engineering, and Technology*. doi:10.4018/IJSSMET.2018070102

Zandia, F., & Tavana, M. (2011). *A fuzzy group multi-criteria enterprise architecture framework selection model*. *Management Information Systems, Lindback Distinguished Chair of Information Systems, La Salle University*. Elsevier.

## **ADDITIONAL READING**

Farhoomand, A. (2004). *Managing (e)business transformation*. Palgrave Macmillan.  
doi:10.1007/978-1-137-08380-7

## **KEY TERMS AND DEFINITIONS**

**Manager:** Business transformation manager.

**Project:** Business transformation project.

# Chapter 7

## A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge, and Behavior Regarding Cyber Security

**Galit Klein**

*Ariel University, Israel*

**Moti Zwilling**

*Ariel University, Israel*

**Dušan Lesjak**

*International School for Social and Business Studies, Slovenia*

### **ABSTRACT**

*With the COVID-19 pandemic, many organizations and institutions moved to e-learning and to e-working from home. With the increase in internet usage, the rate of cyber-attacks have also increased, and this was followed by the request for more cyber security behaviors from employees and students. In the current study, the authors explore the connection between cyber security awareness, cyber knowledge, and cyber security behavior. The authors measured the behaviors among students in two similar countries: Israel and Slovenia. Results show that students felt they had adequate awareness on cyber threat but apply only a few protective measures to protect their devices, usually relatively common and simple ones. The study findings also show that awareness to cyber threats mediate the connection between knowledge and protection behaviors, but only in the case that the knowledge is specific with regard to IT protection courses. Results, implications, and recommendations for effective cyber security training programs for organizations and academic institutions are presented and discussed.*

DOI: 10.4018/978-1-7998-4285-9.ch007

## INTRODUCTION

As the usage of internet increases, cyber security became one of the main concern for private individuals, companies and governments. Cyber threats include various malwares and cybercrime activities, such as the usage of trojan horses, worms, ransomware and spyware malware to perform attacks, collect information, bypass an unauthorised access to data assets and other kinds of hateful behaviours (Srinivas, Kumar Das & Kumar, 2019). These malicious attacks harming and causing disruption to business operations, financial loss but also reduce the trust between computer users and their companies services. In order to response to this problem government legislated several regulation that are aim to protect private and public sectors from crime behaviours, such as the 1996 Health Insurance Portability and Accountability Act (HIPAA) or the Federal Information Security Management Act (FISMA) (Srinivas et al., 2019).

Legislating regulation is just one solution, among others, that increases awareness to cyber hazards among others, including education and training programs. While education process and interactions is not considered as a new idea and has been introduced for several years (Dunn, 2012)<sup>1</sup> the new pandemic accelerated this process. Today, more than ever, children, students, teachers and lecturers are learning through the internets. According to Dunn (2012) the ability to access into unlimited amounts of data, cause them to expand their learning and knowledge horizons, but also to add to their dynamic educational experiences (Dunn, 2012). With that, moving to e-learning environment and relying on cyber technologies which are improved rapidly (from a technological perspective) had yield an increasingly difficult challenges to protect the users from malicious activities and cyber-attacks. As the potential for cyber-attacks became lucid, researches (e.g. Al-Janabi, S., & Al-Shourbaji, 2016; McDaniel, 2013) argue that educational institution should apply cyber awareness programs for cyber protection and cyber security methods. Awareness programs should provide cyber awareness program, cyber knowledge and cyber security active training for users and employees. The programs should be instrumental in developing and spreading security awareness among cyber users, employing proper physical access controls, obeying the security policies and rules as laid down by the institution and the firm in order to achieve the best security outcomes (McDaniel, 2013).

Several cyber security awareness training programs had been presented in the literature affiliated with the awareness program itself (Shaw et al., 2009). Other studies (Lehto, 2015) emphasized the need to understand the factors that motivate or suppress cyber hazard awareness among users. In the current study we will address another angel, and try to reveal *if there is a connection between cyber knowledge, cyber awareness and cyber security behaviours*. To measure these connections we conducted a comparison study in which we compare collected data by Israeli and

Slovenian students from the department of Economics & Business Administration, in both countries. The data was defined by the following variables: cyber knowledge, cyber behaviour and the cyber security awareness. Implications and results are farther discussed.

## **LITERATURE REVIEW**

### **Cyber Security Risks and Solutions**

Since the end of the 20<sup>th</sup> century cyber online transactions has become integral part of our life. As cyber usage becomes more and more prominent, amongst individuals with different levels of knowledge of information technology (IT), there also has been an escalating in the number of cyber-attacks. Cyber threats range from a simple attack, such as spam mails, to a more complicated attacks, such as those conducted by organized cyber-crime groups that use malicious software to steal, corrupt and destroy data stored in people's and organization's devices (Lehto, 2015). For example, in the year 2016, 65% of the UK large businesses reported of being a victim to viruses, spyware or malware. Since this number is only represent companies that reported on a cyber-attack incident we can only assume that the "real" number is much higher (Sharf, 2016). Form individual perspective scholars found that the rate of cyber bullying victimization increased from 18.8% in 2007 to 27.32% in 2010 (Hinduja and Patchin, 2014). Similarly Rek and Milanovski, (2017) reported that most of the secondary school pupils admit to share personal information regarding their life in the networks without having enough aware of the potential exploitation of this information.

According to Srinivas et al. (2019) cyber-attacks are various and include different behaviors that are aimed for different instruments, that have some connection to the networkers. For example attacks can be executed by an implementation of viruses and various attack methodologies, such as the usage of infected programs; phishing attacks, which aim to steal important information from the users, such as information regarding their banking accounts; Planting Trojan horses, which are composed of lines of codes that are directed to execute some harmful function in the users' cyber instruments or portable devices; Worms, which are defined as programs that are actively seek for more instruments in the network to infect them and delete and or utilize important data and files in the system; Spyware which is defined as a program that tracks after crucial and confidential information; Ransomware and Crypto-Ransomwares, which are malwares that prevents or limits the users from accessing their system until a requested payment of money, mostly in cryptocurrency, will be delivered by them (i.e. a ransom money); Unauthorized access, that includes

the access by unauthorized person in order to grab crucial information from the system, including important credentials and passwords. These example are part of a wide range of malicious behaviors that evolve as the usage of computers and the networks increase. Unfortunately this is an arms race where companies that act to produce cyber defense tools against cyber-attacks find themselves facing with new techniques of attacks by new malicious tools which are managed and operated by malicious attackers (adversary).

Understating that the increase in the cyber usage escalate the problem of cyber-attacks leads to requirement for solutions in manner of cyber security standards. Cyber security standards are defined as methods and processes that are implemented and used in an attempt to protect the users or the organizations' cyber environments (Srinivas et al., 2019). These policies and affiliated operational mitigation programs are aimed to reduce and minimize various attacks, and are diversified from macro to micro solutions: from national strategies to individual education programs. The national strategy to secure the cyberspace include among others: legislation of laws, rules and regulations. For example The National Strategy to Secure Cyberspace in the USA is aiming for prevention cyber-attacks against America's critical infrastructures, reducing national system's and infrastructure's vulnerabilities to cyber-attacks and the derived damage as well as the recovery time reduction from cyber-attacks that do acquire<sup>2</sup>. Many companies and individuals search for solutions that can defend themselves from cybersecurity incidents. The solutions are aim to detect and identify cybersecurity incidents before they occur, respond and defend the company's assets or individual's devices from the attack or minimize the damage if the attack already carried out and perform recovery activities in their networks and their systems from the cybersecurity incidents when the attack is detected by IDS (Intrusion Detection Systems) applications. These response categories to cyber threats can also be divided into the following technical solutions: I) Anti-viruses programs or cyber security requirements or procedures, and II) to non-technical solutions, such as cyber security awareness training programs.

## **Security Awareness Training and Educating Programs**

Cyber security behaviors' mapping indicated that individuals are often perform minimum efforts to protect themselves from cyber-attacks. For example, Lukanovič (2017) reported that 40% of the responders in his study argue that they did not know how to install or did not install a protection software on their Internet-connected devices (computer, phone, etc.). The author also report that around 85% of the correspondents admit their computer was infected by virus. While in this study the responders indicated on high percentage of cybersecurity incidents, around one-half of the responders felt that they hold enough information to protect themselves



from the misuse of their personal data. These examples highlight the gap between knowledge and behaviors.

One of the approaches to mitigate cyber security incidents is to develop specific programs that aim to increase the awareness among cyber and network users in an effective way (Kumaraguru et al., 2007). According to the Information Security forum (ISF), security awareness is considered as a “continual process of learning by which, trainees realize the importance of information security issue, the security level required by the organization, and individuals’ security duties”<sup>3</sup>. This definition emphasizes the role of learning and the length of lessons taken by the trainees, however Al-Daeef, Basir and Saudi (2017) also highlight the role of behaviors that should also be taken into account. The authors argue that learning should be considered as the first phase, that must be accompanied by a change in the behaviors of the users. Thus they defined cyber security awareness as “the security knowledge that has been gradually acquired through a continuous and catchy training manner to influence trainees’ behavior” (p. 6).

In addition, Dodge (2007), highlights the need for cyber awareness training in organizations since, in his opinion, most of the recent cyber breaches were caused due to direct involvement of human factors, such as negligent technology operations or inside intruders. Dodge (2007) and Shaw et al., (2009) also suggested that many users are unaware of the cyber risks that council in usage of applications, during the delivery process of information by social networks or that are concealed during the process of web pages surfing along the internet. The scholars suggested that hackers, either individual or organized cyber-crime groups, usually search for the vulnerable squad related to information and networks security which the user is not aware of. The squad is often based on an application, known as a “Software application Bug”, (also known as CVE [Common Vulnerability Exploit] or due to a security breach which was created by the users themselves unintentionally. Therefore, increasing the users’ awareness to these vulnerabilities can increase their readiness to protect the networks from cyber-hazards and cybersecurity incidents.

Over the years scholars explored types of cyber security awareness programs that suggested to students as well as to organizations. For example, Abawajy (2014) explored user preferences regarding cyber security awareness training techniques and methods. Abawajy indicated that cyber education/training can be categorized into different methods depending on delivery technique which contains: online training, contextual training and embedded training. Pawlowski & Yoonhyuk (2015) examined students’ regard to cyber security threats. The scholars succeeded in identifying 23 concepts that are forming the cyber security understandings. Based on their findings they suggested that cyber security courses most concern as a problem centered, during which the course should include case studies that are tailored to students’ level of awareness. In addition they advise to use this taxonomy and alter

security issues from higher-level courses to intermediate and lower ones. Son et al. (2015) also suggested to make differentiation between the programs suggested by the institution. With that, they advise that the program must include some practice plans, such as security labs, to strength the success of the programs.

Although several cyber awareness program are recommended for education institutions there are still existing questions regarding the effect of cyber awareness programs and cyber security behaviors. Dodge (2007) found that phishing scam victims level has dropped among students who were exposed to a program illustrating phishing attacks. McCrohan, Engel., & Harvey (2010) examined the awareness of users to passwords and evaluated ways of securing computers pre and post cyber security training and found that the cyber education/training internet experience created a change in their behaviors. As such the scholars emphasized the need for an appropriate security practices that will change the online behavior in their day-to-day practice. Eminağaoğlu, Uçar., & Eren (2009), showed that awareness campaigns can play as a positive effect in reducing cyber risks among individuals. The authors found that the level of exposure and the extension of training program have forced students to use complex passwords for their own computer during time. The authors argue that providing security awareness to individuals influence the complete reaction toward information security management among employees and individuals. However, changing the password is one of the approaches students can apply to protect their devices. There are many additional means that may not be trivial for the laymen users. In the current study we postulate that *cyber knowledge increases cyber awareness to cyber security problems*. We also intend to explore whether *cyber knowledge will increase cyber awareness to hazard, which will effect the protection behaviors* that the students will carry out to protect their devices from being infected by malicious software. Since Israel and Slovenia have similar GDP values as both considered as developed countries the behaviors of students in both countries was compared.

## **METHODOLOGY**

### **Sample**

A paper-based survey was distributed, among BA students from the department of Economics and Business Administration at Ariel University in Israel (n= 81), and BA and MBA students at the International School for Social and Business Studies from Celje, in Slovenia (n=35). The subjects were located through convenience sampling. Overall the sample included 116 subjects who participated in the survey.

Table 1 exhibits the demographic characteristic of the total sample and for each of the respondents within the countries

*Table 1. demographic Characteristics of the sample*

<b>Characteristics of participants</b>	<b>Israel</b>		<b>Slovenia</b>		<b>Total</b>	
	<b>#</b>	<b>%</b>	<b>#</b>	<b>%</b>	<b>#</b>	<b>%</b>
<b>Total</b>	81	100	35	100	116	100
<b>Male</b>	43	53.1	10	28.6	53	45.7
<b>Female</b>	38	46.9	25	71.4	63	54.3
<b>Type of students</b>						
<b>Bachelor</b>	67	82.7	11	31.4	78	67.2
<b>Master</b>	10	12.3	24	68.6	34	29.3
<b>Ph.D.</b>	4	4.9			4	3.4
<b>Type of study</b>						
<b>Part-time</b>	24	29.6	9	25.7	33	28.4
<b>Full-time</b>	57	70.4	26	74.3	83	71.6
<b>Study field</b>						
<b>Economics</b>	33	40.7	15	42.9	48	41.4
<b>Business &amp; Management</b>	24	29.6	20	57.1	44	37.9
<b>ICT &amp; Logistics &amp; Other</b>	24	29.6			24	20.7

Around 46% of the responders were male, but most of them came from Israel, while the number of females was much higher in the Slovenian group compared to the number of the males in Slovenia. The majority of the students were full-time students, both in Israel and in Slovenia. Most of the Israeli responders, majored in economics, while the Slovenian responders were divided between economic and business administration.

## Instruments

### Survey Study

To measure our assumptions we developed a new questionnaire that aimed to test familiarity of the subjects to cyber in general as well as to test the level of awareness to cyber security risks<sup>4</sup> specifically. Table 2 present the questions and scale that measured the different variables. Each respondent was being asked about former

knowledge in cyber, internet usage and cyber security experience. The classification to different categories was conducted according to the level of respondent's knowledge to cyber security and cyber threats (*Knowledge*), their cyber security awareness (*Awareness*), their familiarity with cyber security incidents, and their attempts to control and prevent cyber-attack (*Behavior*).

## Fuzzy Logic Study

A fuzzy testing concept, or more specifically, fuzzy hypothesis testing is considered as a verification-based method in machine learning. The fuzzy hypothesis test is used to determine the truth (or falsity) of a proposed hypothesis. The fuzzy model is executed on the fuzzy data and produce a value on  $[0,1]$  indicates on the degree to which the hypothesis is considered as valid for a given sample data.

A fuzzy set, as defined by Zadeh (1965) is composed of several elements: Inputs that are connected to outputs through member functions. The model used is comprised of three steps as described by Zwilling (2020):

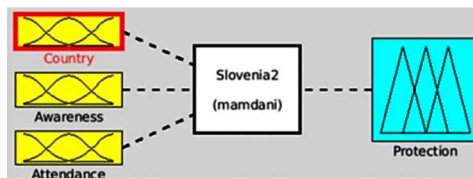
- “The first step of Fuzzification is executed as the transformation of numerical values into ordinary language, if necessary. For example, the inputs hold linguistic values such as Low, Medium, and High. Each variable usually contains three to seven terms (Attributes).”
- “The second step is expressed in a “Fuzzy Inference” which defines the system behavior by means of the rules such as <if>, <and>, <then>, <with>. The conditional clauses create rules, which evaluate the input variables. These conditional clauses were designed as follows:
- <if>  $x_1$  is value  $x'_1$  <and>  $x'_2$  is value  $x'_2$  ... <and>  $x'_N$  is value  $x'_N$  <then>  $y_1$  is value  $y'_1$  <WITH> probability  $s$ , where  $x_i$  are inputs,  $y_j$  are outputs,  $x'_i$  are values of inputs,  $y'_j$  are values of outputs and  $s$  is a degree of support”.
- “The third step (“Defuzzification”) is expressed by transformation of linguistic values into numerical ones, if necessary. The outputs also use linguistic values and various types and shapes of membership functions”.

“The fuzzy logic was described by the following terminology: A fuzzy set  $A$  is defined as  $(U, \mu_A)$ , where  $U$  is the relevant universal set and  $\mu_A: U \rightarrow [0,1]$  is a membership function, which assigns each element from  $U$  to the fuzzy set  $A$ . The membership of the element  $x \in U$  of a fuzzy set  $A$  is indicated as  $\mu_A(x)$ . We define  $F(U)$  as the set of all fuzzy set. Then the “classical” set  $A$  is the fuzzy set where:  $\mu_A: U \rightarrow \{0, 1\}$ . Thus  $x \notin A$ ,  $\mu_A(x) = 0$  and  $x \in A$ ,  $\mu_A(x) = 1$ . Let  $U_i$ ,  $i = 1, 2, \dots, n$ , be universals. Then the fuzzy relation  $R$  on  $U = U_1 \times U_2 \times \dots \times U_n$  is a fuzzy set  $R$  on the universal  $U$ ”.

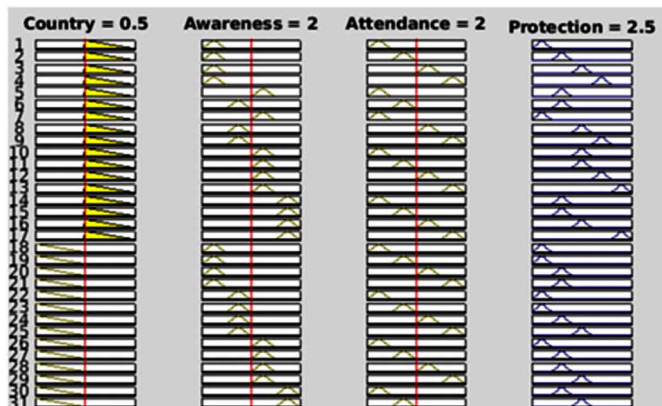
A fuzzy model based questionnaire data for hypothesis testing validation was used for the first study (Figure 1). The Sav data was transformed into a csv file. In order to create the mapping functions for each fuzzy hypothesis, the fuzzy sets corresponding to “Country”, “Awareness”, “Attendance” and “Protection” have been determined. The member functions were chosen with a triangle function and the set of rules was set according to the appropriate scale (Figure 2). Finally the output was analyzed by the following code:

1. `fis = ('Slovenia2.fis')`
2. `input = readmatrix('Slo.csv')`
3. `Output = evalfis(fis, input)`

*Figure 1. Fuzzy Logic Model Design.*



*Figure 2. Fuzzy Logic Rules*



## RESULTS

### Descriptive Analysis

Descriptive analysis was initially conducted to capture the amount of awareness, knowledge and behaviors toward cyber-attacks. The results of the mean and standard deviation scores, for the total countries and in each country alone, are presented in the Table 2.

Table 2. Descriptive statistics

Variables' name	Description/Question	Descriptive Statistics		IL (n=89)	SI (n=35)	T/ Chi²
		Mean (Sd.)	Yes *	Mean (Sd.)	Mean (Sd.)	
<b>Awareness</b>	Are you familiar with the term cyber security (1- no knowledge to 4- very good)	2.44 (.81)		2.39 (.84)	2.57 (.60)	T= 1.10 p>0.05
<b>Familiarity</b>	Familiarity of different sources – total score (range from 0- through 9 sources)	3.98 (.81)		3.56 (1.91)	4.85 (2.36)	T= 3.83**
<b>IT past</b>	Attendance in IT security training in the past		21.8%	23.6	17.1	X²= .647
<b>IT future</b>	Would like to attend in IT security training, (1- definitely not to 5- definitely yes)	3.74 (.97)		3.65 (1.0)	3.97 (.85)	T=1.66¹
<b>Threats</b>	The main cyber security threats are- (1- strongly disagrees to 5- strongly disagree)	4.04 (.87)		4.06 (.88)	4.00 (.86)	T= -.35, p>0.05
<b>Effect of Education on awareness</b>	The extend in which the current education influenced their cyber-security awareness (1- definitely not affected to 5- strongly affected)	3.22 (.98)		3.15 (.90)	3.40 (.95)	T=1.29, p>0.05
<b>Recognition</b>	I usually recognize and know the differences between http and https protocol		51.6%	52.8%	48.6%	X²= .18
<b>Provide</b>	Total sum of the amount of information that the responders provide in the web (1-strongly disagree to 5- strongly disagree)	2.42 (.93)		2.54 (.90)	2.13 (.95)	T= -2.19*
<b>Computer knowledge</b>	Self-evaluation of skills and knowledge in using computer application (range from 1-no skills to 5- very high skills)	3.22 (.67)		3.25 (.70)	3.16 (.60)	T=.62, p>0.05
<b>Behavioral</b>	I know how to behave in case of cyber-attack (1- definitely no to 5-rather yes.)	3.10 (1.19)		3.07 (1.19)	3.17 (1.20)	T=.66, p>0.05
<b>Choice</b>	Is the use of technology products coming from your desire or by coercion (1-definitely by coercion to 5- definitely by choice)	3.22 (1.00)		3.40 (.98)	3.11 (1.02)	T=1.46, p>0.05
<b>Protection</b>	Sum of the score in the usage that the responders make to protect their instrument (ranged from 0 to 11)	4.09 (2.23)		3.50 (1.87)	5.60 (2.39)	T=5.16***
<b>Length</b>	The average length of your standard password (minimum 0 to maximum -14)	9.43 (5.27)		8.99 (5.72)	10.49 (3.85)	T=1.41, p>0.05
<b>Password</b>	Do you use the same password for different portals, system and application		54.8%	63	34.3	X²=8.31**
<b>Finish</b>	Sum of the activities that the responders are acting when finish working on the computer (range from 0 to 4)	1.37 (.73)		1.32 (.67)	1.48 (.88)	T=1.08, p>0.05

Notes: ¹ = standard deviation appears in the parentheses; ²= yes represent the percentage of people that were argue they agree with the sentence. It is relevant for dichotomy questions (i.e. yes/no questions). \* for the categorical variables, the results indicate the percentage of respondents who agree with the item from the corresponding country.  
¹p< 0.10; \*p<0.05; \*\*p<0.01

The result indicate that responders have medium level of awareness to cyber security (M=2.44), but it was a somewhat higher among the Slovenian students compared to the Israeli values (M=2.59 and M=2.39 respectively). The overall knowledge ranged between lower to low knowledge, across the different knowledge measurement. The responders felt they have medium level amount of computer knowledge (M=3.22) and half of them argue that they know the differences between

http and https protocol. However, when we asked about the degree of familiarity to different cyber protection tools out of optional 9 tools they were familiar in average with 4. When asking them about their involvement and participation in IT security courses, only 21% answered that they participate in such courses. Analysis of their behaviors regarding cyber security reveals that the subjects made some efforts to protect their tools, but it wasn't ultimate. The responders defend their password with longer length (around 9 characters), and felt that they know how to behave in the case of cyber-attack ( $M=3.22$ ). On the other hand only few protection activities when finishing working on their computers were reported. Therefore it appears that the subjects are aware of the hazards in the web and make some attempts to defend themselves from cybersecurity incidents.

### **The Connection Between Cyber Security Knowledge, Cyber Security Awareness and Cyber Security Behaviors**

Our main concern was to assess the factors that affect cyber security behaviors and whether cyber security awareness mediate the connection between cyber knowledge and cyber security behaviors. Particularly we wanted to explore and understand if attendance in cyber security courses are related to more hazard behaviors. To analyze our assumption we conducted a path analysis by the Amos<sup>TM</sup> software with two independent variables: *computer knowledge*, indicating general knowledge regarding the usage of computer (model 1-2) and *attendance in IT courses in the past* (model 3-4). We also measured independent behaviors with two variables: *Protection*, representing the amount of measurement that the subject use to protect their devices from cyber security incidents. In addition we measured the *provide* variable which represent the amount of information that the responders are willing to share on the web. Higher result indicate on less secure behavior. Lastly, we control the students country. Table 3 exhibits the result of the path analysis.

We first measured general computer knowledge as our independent variables. We wondered if computer knowledge raise awareness which lead to more cyber *protection* behaviors. The results, as appear in table 3 model 1, indicate that the model fit was not adequate:  $\chi^2= 16.14$ ,  $df=2$   $p<0.05$ ,  $TLI=.36$ ,  $CFI=.79$ ,  $RMSEA=.24$ .

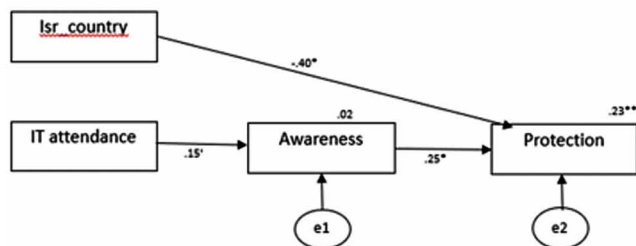
In the second analysis we measured whether computer knowledge is connected to higher readiness to provide information through the web via the mediation of cyber awareness. The results, as appear in table 3 model 1, indicate that the model fit was not adequate:  $\chi^2= 16.08$ ,  $df=2$   $p<0.05$ ,  $TLI= -.11$ ,  $CFI=.63$ ,  $RMSEA=.24$ . Based on the two analyses we can conclude that the awareness for cyber security does not affect the connection between computer knowledge and cyber security behaviours<sup>5</sup>.

### A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge

Next we measured the connection of attending in IT security courses with cyber security behaviors via the mediation effect of the awareness to cyber security. Table 3 exhibits model 3 fit results. The model is illustrated in Figure 3.

*Figure 3. Connection between attendance in IT course in the past, cyber security awareness and protection behaviors.*

*Boldface arrows indicate structural component. e = error.*



Overall the fit of model 3 measurement was adequate:  $\chi^2=1.28$ ,  $df=2$ ,  $TLI=1.07$ ,  $CFI=1.00$ ,  $RMSEA=.00$ . The result shows that attendance in IT security in the past was close to significant with the awareness to cyber security ( $\beta=.15$ ,  $p<0.10$ ). Awareness was also connected with cyber protection behaviours ( $\beta=.25$ ,  $p<0.01$ ). In addition the bootstrap results showed that the indirect effect of attendance in IT security courses and protection behaviours via to cyber security awareness was significant (.11, B-CCI=.07-.018,  $p<0.00$ ). We also found that the connection between the country and protection behaviour was significant ( $\beta=-.40$ ,  $p<0.01$ ) indicating that the Slovenian students make more cyber protections behaviors.

Lastly, we analysed if the connection between attendance in IT security course in the past is connected to the readiness to provide information, via the mediation

*Table 3. The results of path analysis*

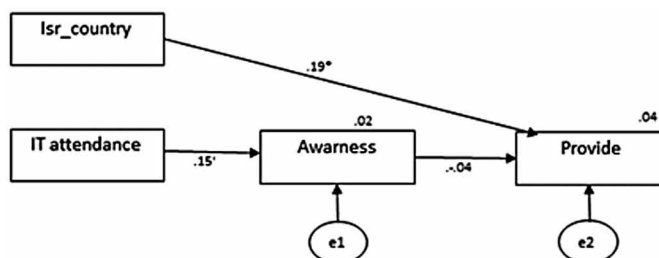
Model	Root	$\chi^2$	df	CFI	NFI	TLI	RMSEA	$\chi^2/df$	95%CI
Model2	Computer knowledge--- awareness---protection	16.14	2	.79	.78	.36	.24	8.07	-.03, .11
Model2	Computer knowledge--- awareness---provide	16.08	2	.63	.63	-.11	.24	8.04	-.08, .02
Model3	<b>IT_attendance</b> --- awareness---protection	1.28	2	1.00	.96	1.07	.00	.64	.07, .18
Model4	<b>IT_attendance</b> --- awareness---provide	1.33	2	1.00	.85	1.65	.00	.69	-.02, -.00



of the cyber security awareness. Figure 4 illustrated the result of path analysis and table 3 exhibits model 4 fit results.

*Figure 4. Connection between attendance in IT course in the past, cyber security awareness and the readiness to provide information in the web.*

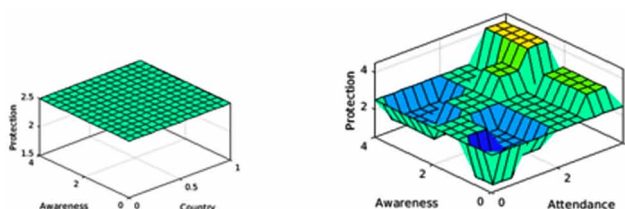
*Boldface arrows indicate structural component. e = error.*



Overall the fit of model 4 measurement was adequate:  $\chi^2=1.33$ ,  $df=2$ ,  $TLI=1.65$ ,  $CFI=1.00$ ,  $RMSEA=.00$ . The result shows that the attendance in IT security in the past was close to significant related with the awareness of cyber security ( $\beta=.15$ ,  $p<0.10$ ), but awareness was not connected with the readiness to provide information on the web ( $\beta=-.04$ ,  $p>0.05$ ). The bootstrap results showed that the indirect effect of attendance in IT security courses and the readiness to provide information via the awareness to cyber security was significant (.09,  $BCCI=-.02-.001$ ,  $p<0.00$ ). We also found that the connection between the country and protection behaviour was significant ( $\beta=.19$ ,  $p<0.05$ ) indicating that the Israeli students have more willingness to provide information through the web networks, indicating on more hazard behaviors.

The results indicate that while general knowledge may not affect the awareness to cyber security or to cyber security behaviors, attending in IT security courses do connect through the mediation of awareness to cyber knowledge.

*Figure 5. Output of Fuzzy Logic Surface related to the fit model*



## **Conformation of the Statistical Model Based on Fuzzy Logic Model**

The fuzzy logic model was found as supporting the study findings. It was found that Slovenian students who attended IT Training courses in the past, have more awareness to cyber threats and their cyber protection behavior is influenced by these two variables, so as they are less exposed to cyber threats than their Israeli students. The Surface output is shown in Figures 5.

## **CONCLUSION**

The current study results show that internet users aware to the term “cyber security”. The majority of the participants in our study felt they had adequate awareness to cyber hazard. However, being aware of a problem does not mean that people are taking means to prevent them from being attacked. Similar to former studies (e.g. Imgraben, Engelbrecht, & Choo, 2014; Rek & Milanovski, 2017) we found that responders made basic and not sufficient activities. For example they used long password, and tried not to provide many information on the web. On the other hand, when they had to do active behaviors, such as conducting different procedures, when finishing using the computer, we found that the readiness to protect their device decrease. As such we can found a gap between active and passive behaviors. Since long password are mostly mandatory by the institution, we recommend that more protection action should be implant as mandatory action, to protect organizational tools.

We also found that awareness mediate the connection between knowledge and cyber security behaviors. With that, the result were somewhat surprising. Cyber hazard awareness did not affect the connection between general cyber knowledge and cyber security behaviors. On the other hand, awareness to cyber hazard mediate the connection between attendance in IT security courses and cyber security behaviors. E.g. students that participants in IT security courses in the past had higher awareness to cyber hazards, therefore conducted more protection action and provide less information on the web. These results is in congruent with former studies that indicated on the role of cyber security education programs in promoting security behaviors (Abawajy, 2014; Al-Daeef et al., 2017; Dodge, 2007). Our research reinforces the previous claims and suggests that one of the reason for the connection between cyber hazard knowledge and cyber protection behavior stems from an awareness to the potential threats in the web.

The connection between knowledge and behaviors can be explained through the Theory of Planned Behavior (TPB) (Fishbein & Ajzen, 2010). According to the theory, intention is the best predictor of planned behavior. Thus if users are aware of

the threat of using computers highly it will motivate them to take multiple action to protect their computer. With that, behavior is also affected from other elements, such as the amount of self-efficacy and controllability. Therefore, when the responders perceived themselves as being able to control the situation by their behaviors, their motivation to take actions increases. This may explain why participants tend to conduct behaviors that were mandatory or required simple computer knowledge, but fail to apply more sophisticated actions to protect their devices from being abused. The results highlight the importance role of cyber security programs to motivate the cyber users for proactive behaviors. It also highlight the need to implant simple protection method that will compensate on the lack of cyber knowledge.

Lastly, we found a connection between the responder's country and the cyber security behaviors. Israeli students tend to make fewer protection behaviors and delivered more information on the web, compare to Slovenian students. One explanation is that Israelis feels that they live in a leading high-tech technology country, which is perceived as a cyber country or cyber nation (Tabansky, 2013). Thus, they may rely on their organization and country to implement protective method that will secure their devices. As individuals, they do not feel that they need to be active in protecting their devices or their organizational assets. Indeed, in his study, Tabansky (2013), describes Israel as a country that continuously strive to develop cyberspace solutions according to the change in opportunity and risks. The systematic perception of Israeli as facing against global cyber threats lead many high-tech firms in Israel to develop cyber defense technologies and invest in cyber protection R&D. Based on this assumption we believe that the more the country is developed in their substantial GDP value and tend to invest in cyber security protections measurements (such as in the case of Israel), the less the citizens make effort to conduct active protection behaviors and rely on outsiders entities to protect them. Therefore, more efforts should be invested in educating and increasing the awareness among the citizens of these countries. The result also indicated that the awareness, knowledge and especially behaviors are effected by cultural defense, values and demographic characteristic. more research should explore effects of psychological factors, such as self-efficacy and national-cultural values (Hofstede, 2001; Klein et al., 2016; Klein & Shtudiner, 2016; Shtudiner et al., 2019) on Internet user behaviors. future studies are needed to explore the active actions from senior management in the evaluated countries and may be considered as a follow up research.

During the last decades, the need to mitigate cyber risks awareness among employees and individuals had increased. Based on the results of the current study we present several recommendation that we believe will assist to increase cyber security awareness and protect personal and organizational devices.

- Organizations should provide risk management framework or internal policies, similar to the framework provided by The National Institute Standards and Technology (NIST)<sup>6</sup>. The cybersecurity framework (CSF) offers free guidance, based on existing regulation, standards, guidelines and practices. This framework intend to aid the organizations to better protect themselves from cyber security threats. E.g. The NIST framework indicate five core functions: Identify, Protect, Detect, Respond and Recover from attacks. However, organizations will need to customize the cybersecurity framework so it will suit the specific threats that firms is more vulnerable to. For example, universities must protect their students' and lecturers' information, while financial institutions face more extreme hazards concerning the privacy and financial threats. Therefore while the framework is general, each firm should tailor a specific protection method according to its aims and processes.
- Organizations, especially educational institutions, need to provide cyber security and cyber protection training courses. These courses duration can be either long (e.g. semester) or given as a workshop and modules (i.e. short ones), similar to the NIST courses<sup>7</sup>. In any case, these courses should be mandatory for the employees and students. We also recommend that students and employees will need to undergo training session on cyber threats and on cyber protection's tools on a regulatory bases, since the world of cyber hazard and cyber protection changes rapidly.
- Organizations need to oblige employees and students to implement methods that will protect organizational and individual devices. For example, to set a password that will be long enough and multi-character, so it will reduce the chance of cracking the password by Brute Force technique, for example. Organization can provide protection tools, such as antivirus software, that will be appealing for implementation by the workers and students.
- In countries that have higher awareness to high-tech and cyber R&D, similarly to Israel, organizations should explore more thoroughly the level of awareness and behaviors of their employees concerning cyber threats and cyber protection. They also should make specific courses that will highlight the gap between the national protection programs and individuals protection behaviors. As we found in our study, Israeli students made less protection behaviors compare to the Slovenian counterparts even though the country is well recognize and perceived by its high-tech development in cyber within High-Tech companies.
- Organization must rely on themselves and obtain a reliable network protection tools. Organization should not rely on the voluntary behaviors of their employees, not yet alone students. As such, they need to determine clear standards, policy and protection measurements. If the organization

has number of branches or cooperation with other institution (resemble to university and research centers) standards and protection methods should be set in all of organizational branches and centers.

To conclude, with the new decade and the Covid-19 pandemic, more organizations and institutions converted to e-learning and working from home procedures (e-working). With the increase usage of internet and internet tools usage, so is the growth in the risk of being a victim to cyber-attack. Future work should focus on exploring how specific training programs based on the study findings improve the level of cyber knowledge and skills. We believe that knowledge is the basic for behavioral changes. Therefore cyber security education can assist in preventing cyber-attacks and reduction of damages caused to networks, databases and network-components due to cyber-attacks performed by organized-cybercrime groups or cybercrime individuals.

## REFERENCES

- Abawajy, J. (2014). User preferences of cyber security awareness delivery methods. *Behavior & Information Technology*, 33(3), 236-247.
- Al-Daeef, M. M., Basir, N., & Saudi, M. M. (2017, July). Security awareness training: A review. In *Proceedings of the World Congress on Engineering (Vol. 1, pp. 5-7)*. Academic Press.
- Al-Janabi, S., & Al-Shourbaji, I. (2016). A study of cyber security awareness in educational environment in the middle east. *Journal of Information & Knowledge Management*, 15(1). doi:10.1142/S0219649216500076
- Dodge, R. C. Jr, Carver, C., & Ferguson, A. J. (2007). Phishing for user security awareness. *Computers & Security*, 26(1), 73–80. doi:10.1016/j.cose.2006.10.009
- Dunn, J. (2012). *The importance of internet access in schools*. Available at <http://www.edudemic.com/2012/12/the-importance-of-internet-access-in-schools/>
- Eminağaoğlu, M., Uçar, E., & Eren, S. (2009). The positive outcomes of information security awareness training in companies – A case study. *Information Security Technical Report*, 14(4), 223–229. doi:10.1016/j.istr.2010.05.002
- Fishbein, M., & Ajzen, I. (2011). *Predicting and Changing Behavior: The reasoned Action Approach*. Psychology Press. doi:10.4324/9780203838020

Hinduja, S., & Patchin, J. W. (2014). *Cyberbullying*. Cyberbullying Research Center. Retrieved September, 7, 2015; retrieved from <https://cyberbullying.org/Cyberbullying-Identification-Prevention-Response.pdf>

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, 2(1), 2307–2319. doi:10.9707/2307-0919.1014

Imgraben, J., Engelbrecht, A., & Choo, K. (2014). Always connected, but are smart mobile users getting more security savvy? A survey of smart mobile device users'. *Behaviour & Information Technology*, 33(12), 1347–1360. doi:10.1080/0144929X.2014.934286

Klein, G. (2016). Trying to make rational decisions while employing intuitive reasoning: A Look at the due-diligence process using the dual-system reasoning model. *International Journal of Entrepreneurship and Innovation Management*, 20(3/4), 214–234. doi:10.1504/IJEIM.2016.077962

Klein, G., & Shtudiner, Z. (2016). Trust in others: Does it affect investment decisions? *Quality & Quantity*, 50(5), 1949–1967. doi:10.1007/11135-015-0245-6

Kumaraguru, P., Acquisti, A., Rhee, Y., & Nunge, E. (2007). Protecting people from phishing: The design and evaluation of an embedded training email system. *Conference on Human Factors in Computing Systems – Proceedings*, 905-914. 10.1145/1240624.1240760

Letho, M. (2015), Cyber Security Competencies - Cyber Security Education and Research in Finnish Universities. *14th European Conference on Cyber Warfare and Security (ECCWS)*, 179-188.

Lukanović, L. (2017). *Računalniška kriminaliteta in varstvo osebnih podatkov: diplomatska naloga*. Available at: [http://www.ediplome.fm-kp.si/Lukanovic\\_Lea\\_20171017.pdf](http://www.ediplome.fm-kp.si/Lukanovic_Lea_20171017.pdf)

McCrohan, K., Engel, K., & Harvey, J. (2010). Influence of awareness and training on cyber security. *Journal of Internet Commerce*, 9(1), 23–41. doi:10.1080/15332861.2010.487415

McDaniel, E. (2013, July). Securing the information and communications technology global supply chain from exploitation: Developing a strategy for education, training, and awareness. In *Proceedings of the Informing Science and Information Technology Education Conference* (pp. 313-324). Informing Science Institute. 10.28945/1813

- Pawlowski, S., & Yoonhyuk, J. (2015). Social representations of cyber security by university students and implications for instructional design. *Journal of Information Systems Education*, 26(4), 281–294.
- Rek, M., & Milanovski, B.K. (2017). *Slovenija, Ljubljana: Fakulteta za medije [izdelava]*. Slovenija, Ljubljana: Univerza v Ljubljani, Arhiv družboslovnih podatkov [distribucija], IDNo: MPSS16.
- Sharf, E. (2016). Information exchanges: regulatory changes to the cyber-security industry after Brexit: Making security awareness training work. *Computer Fraud & Security*, 2016(7), 9–12. doi:10.1016/S1361-3723(16)30052-5
- Shaw, R. S., Chen, C. C., Harris, A. L., & Huang, H.-J. (2009). The impact of information richness on information security awareness training effectiveness. *Computers & Education*, 52(1), 92–100. doi:10.1016/j.compedu.2008.06.011
- Shtudiner, Z., Klein, G., Zwilling, M., & Kantor, J. (2019). The value of souvenirs: Endowment effect and religion. *Annals of Tourism Research*, 74, 17–32. doi:10.1016/j.annals.2018.10.003
- Srinivas, J., Das, A. K., & Kumar, N. (2019). Government regulations in cyber security: Framework, standards and recommendations. *Future Generation Computer Systems*, 92, 178–188. doi:10.1016/j.future.2018.09.063
- Tabansky, L. (2013). Critical Infrastructure Protection Policy: The Israeli Experience. *Journal of Information Warfare*, 12(3), 78–86.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. doi:10.1016/S0019-9958(65)90241-X
- Zwilling, M., Levy, S., Gvili, Y., & Dostal, P. (2020). Machine learning as an effective paradigm for persuasive message design. *Quality & Quantity*, 54(3), 1–23. doi:10.1007/11135-020-00972-0

## ENDNOTES

- <sup>1</sup> Acknowledgment: The research was funded by the Ariel Cyber Innovation Center.
- <sup>2</sup> <https://georgewbush-whitehouse.archives.gov/pcipb/>
- <sup>3</sup> Information Security Forum (ISF): The Standard of Good Practice for Information Security. Security Standards. 2007.

***A Comparative Study in Israel and Slovenia Regarding the Awareness, Knowledge***


- <sup>4</sup> Detail information regarding the instruments can be found in Zwilling, M., Klein, G., Lesjak, D., Wiechetek, L., Faith, C., Hamdullah, N.B. (2020). Cyber security knowledge, cyber threat awareness and cyber user behavior: A comparative behavioral study of Israel, Poland, Slovenia and Turkey. *Journal of Computer Information System*. 1-16 <https://www.tandfonline.com/eprint/PX76PBVGPDYKV8VNYZYS/full?target=10.1080/08874417.2020.1712269>
- <sup>5</sup> We also measured a direct connection between computer knowledge and cyber security behaviours. The model fit was still not adequate ( $\chi^2= 21.56$ ,  $df=2$   $p<0.05$ ,  $\chi^2= 22.31$ ,  $df=2$   $p<0.05$ ).
- <sup>6</sup> <https://www.nist.gov/cyberframework>
- <sup>7</sup> <https://www.nist.gov/cyberframework/online-learning>



## Chapter 8

# Developing Cyber Security Competences Through Simulation-Based Learning

**Bistra Konstantinova Vassileva**

 <https://orcid.org/0000-0002-5976-6807>  
*University of Economics, Varna, Bulgaria*

### ABSTRACT

*The importance of cyber security competences is growing both in practice and in academia during the last few years. This chapter provides a current overview of the existing body of the literature in the field of simulation-based learning and the key cyber security issues. The author's primary goal is to develop a methodological business-oriented and evidence-based learning framework which will provide students or trainees with the opportunity to develop practical skills in the field of cyber security issues through a virtual business simulator. The overall intention is to provide a coherent framework that makes use of active-based learning and gamification to support the active participation of students or trainees. To meet these goals, the Reference Framework for Applied Competences (REFRAC) is applied. Taking into account that in 2040 ICT and internet will be 'culturally invisible', cyber security competences will be a must for everyone. They will be critical both for personal and companies' survival in the turbulent and highly competitive digital environment.*

DOI: 10.4018/978-1-7998-4285-9.ch008

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

## **INTRODUCTION**

The importance of cyber security competences is growing both in practice and in academia during the last few years. This chapter provides a current overview of the existing body of the literature in the field of simulation-based learning and the key cyber security issues. The author's primary goal is to develop a methodological business-oriented and evidence-based learning framework which will provide students or trainees the opportunity to develop practical skills in the field of cyber security issues through a virtual business simulator. The overall intention is to provide a coherent framework that makes use of active-based learning and gamification to support active participation of students or trainees in the learning process. To meet these goals, the Reference Framework for Applied Competences (REFRAC) is applied. Taking into account that in 2040 ICT and internet will be 'culturally invisible' (Manyika et al., 2015) cyber security competences will be a must for everyone. They will be critical both for personal and companies' survival in the turbulent and highly competitive digital environment. Research questions driving this chapter are as follows: 1/ to identify the key topics of cyber security which should be taken as mandatory topics during the training sessions; 2/ to evaluate the possibilities of simulation-based learning to be applied for cyber security issues, and 3/ to propose a methodological framework of simulation-based learning environment aimed at cyber security skills development.

## **BACKGROUND**

This chapter begins with outline of the importance of cyber security issues, cyber security education and experience-based learning approach. The author's primary goal is to develop a methodological business-oriented and evidence-based learning environment which will provide students the opportunity to experience different professional skills, incl. cyber security competences. The overall intention is to offer a coherent framework that is student-oriented and makes use of active-based learning to encourage student active participation. A survey among students was conducted to support the identification of critical cyber security competences to be used in the background layer of the Reference Framework for Applied Competences (REFRAC).

Worldwide spending on information security products and services is estimated to reach over \$124 billion in 2019 (RSAC, 2019). Cyber security budgets have been on the rise for the past several years, increasing by 141% from 2010 to 2018. These numbers show the raising concern to the new challenges to legitimate businesses caused by the increasing activities of the cyber criminals. Cyber security is becoming a key business enabler and a vital tool to protect competitive advantage of companies

(Buffomante, 2020:1). According to the World Economic Forum (WEF), the rising cyber interdependence of infrastructure networks is one of the world's top risk drivers. The WEF 2017 Global Risks Report found that cyberattacks, software glitches, and other factors could spark systemic failures that “cascade across networks and affect society in unanticipated ways” (WEF, 2017:7).

## **MAIN FOCUS OF THE CHAPTER**

### **The Challenges of Cyber Security Landscape**

Security is not a new concept but it is of vital importance nowadays when significant security incidents are a regular occurrence. Globalization and advances in technology have driven unprecedented increases in innovation, competitiveness, and economic growth. Critical infrastructure has become dependent on these enabling technologies for increased efficiency and new capabilities (NIST, 2014).

The key findings from The Global State of Information Security Survey 2018 done by KPMG (Castelli, Gabriel, Yates and Booth, 2020) show that massive cybersecurity breaches have become almost commonplace, regularly grabbing headlines that alarm consumers and leaders. Such strong dependence on ICT raises the following question to the academics in the field of economics and business: How well are we prepared to teach our students to be prepared to work in complex cyber threat landscape and highly competitive marketplace?

During the past decade the following four major drivers setting directions to security decision makers in organizations have been identified (Dias et al., 2017): government and industry-sector-specific regulations; standards and best practice models for IT security; business risks and security requirements of the business network that an organization has or wants to join; urgency to invent opportunities in the midst of security breakdowns that incur monetary damage, corporate liability, and loss of credibility. The additional reasons for paying attention on cyber security issues are rooted in the following market trends:

- increasing frequency and cost of security breaches;
- the speed of business activities in the post-economic-crisis world, incl. new product launches, M&A, market expansion, and introductions of new technologies;
- global accessibility of data due to high penetration rates and proliferation of mobile computing (use of internet, smartphones and tablets in combination with BYOD);
- continually complicated ecosystem of digitally connected entities;

- cloud-based services, and third party data management and storage;
- open technology systems.

Therefore, cyber security is growing in complexity every day and requires continual refinement of the workforce's capacity for both skill and strategy (Tipton, 2014). The main results of research on cyber security are provided in Table 1.

The economic aspects of cyber security, addressing the specific risks and vulnerabilities of information technologies were investigated as well (Herzog, 2003; Baryshnikov, 2012; Bissell, Lasalle and Dal Cin, 2020). They became increasingly scrutinised, especially given the relatively open, standards driven character of most communication and data processing protocols (FireEye, 2020). According to Anderson (2008) the business impact of cyber security breaches can be classified into four broad categories. The first category is financial impact which results in loss of sales, loss of tangible assets, unforeseen costs, legal liabilities, and depressed share price. The operational impact includes loss of management control, loss of competitiveness, breach of operating standards. Customer-related impact stems from loss of customers or clients, loss of confidence, reputational damage, and delayed deliveries. Employee-related impact mainly results in reduction in staff morale and/or productivity. The above mentioned reasons explain why training on cyber security in the field of economics is needed. This conclusion is confirmed by the results of EY Global Information Security Survey 2018-19 and its implications for digital economy. Effective cyber security becomes increasingly complex to deliver which requires strategic business transformation especially toward the business ecosystem. Taking into account that many systems fail because their designers protect the wrong things, or protect the right things but in the wrong way (Anderson, 2008), we could assume that for an organization to be able to effectively manage the risks in its ecosystem, it needs to clearly define its limits, main components and critical business process which are vulnerable to security breaches. This corresponds to the more horizontal, strategic approach to cyber security in contrast to the vertical, technical approach.

Recently the focus of cyber security research shifted to the role of people (wetware, liveware, meatware, humanware) (Harley, 2010) and especially to millennials-focused talent management presented in 2018 Deloitte-NASCIO Cybersecurity Study. A group of researchers and academics (Guzdial, 2015; Jurse and Mulej, 2011) emphasised the necessity to re-target the education from knowledge transfer to development of a mental models/mindsets which (1) make users vulnerable to cyber attacks, and (2) make management resistant to addressing cyber security issues (Harley, 2010).

*Table 1. Attitudes toward cyber security: main reports and research findings*

Title	Sample Size	Source	Most Important Findings
All hands on deck: Key cyber security considerations for 2020		<a href="https://home.kpmg/xx/en/home/insights/2020/03/key-cyber-security-considerations-for-2020.html">https://home.kpmg/xx/en/home/insights/2020/03/key-cyber-security-considerations-for-2020.html</a>	<p>Six key cyber considerations:</p> <ul style="list-style-type: none"> <li>§ Aligning business goals with security needs.</li> <li>§ Digital trust and consumer authentication.</li> <li>§ The evolving security team.</li> <li>§ The next wave of regulation.</li> <li>§ Cloud transformation and resilience.</li> <li>§ Automating the security function.</li> </ul>
2019 HIMSS Cybersecurity Survey	166	<a href="https://www.himss.org/himss-cybersecurity-survey">https://www.himss.org/himss-cybersecurity-survey</a>	<ul style="list-style-type: none"> <li>§ A pattern of cybersecurity threats and experiences is discernable across US healthcare organizations.</li> <li>§ Many positive advances are occurring in healthcare cybersecurity practices.</li> <li>§ Complacency with cybersecurity practices can put cybersecurity programs at risk.</li> <li>§ Notable cybersecurity gaps exist in key areas of the healthcare ecosystem.</li> </ul>
The Global State of Information Security Survey 2018	9500	<a href="https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey/strengthening-digital-society-against-cyber-shocks.html">https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey/strengthening-digital-society-against-cyber-shocks.html</a>	<ul style="list-style-type: none"> <li>§ 39% say they are very confident in their cyberattack attribution capabilities.</li> <li>§ The frequency of organizations possessing an overall cybersecurity strategy is particularly high in Japan (72%) and Malaysia (74%).</li> <li>§ Only 44% of the respondents say their corporate boards actively participate in their companies' overall security strategy.</li> <li>§ 34% say their organizations plan to assess IoT risks across the business ecosystem.</li> <li>§ Only half of respondents say their organizations conduct background checks.</li> <li>§ Only 58% of respondents say they formally collaborate with others in their industry, including competitors, to improve security and reduce the potential for future risks.</li> </ul>
The Global Risks Report 2017 12 <sup>th</sup> Edition	745	<a href="http://www3.weforum.org/docs/GRR17_Report_web.pdf">http://www3.weforum.org/docs/GRR17_Report_web.pdf</a>	<ul style="list-style-type: none"> <li>§ The next global challenge: facing up to the importance of identity and community.</li> <li>§ The risks associated with AI are considered the potential risks associated with letting greater decision-making powers move from humans to AI programs, as well as the debate about whether and how to prepare for the possible development of machines with greater general intelligence than humans.</li> <li>§ Rising cyber dependency is determined as number 4 within the Top 5 Trends that determine global developments.</li> </ul>
2015 HIMSS Cybersecurity Survey	297	<a href="https://www.himss.org/2015-cybersecurity-survey">https://www.himss.org/2015-cybersecurity-survey</a>	<ul style="list-style-type: none"> <li>§ The most important problem is the breach of patient information.</li> <li>§ The respondents' organizations use an average of 11 different technologies to secure their environments.</li> <li>§ 50% of respondents' organizations have hired a full-time professional, such as a Chief Information Security Officer (CISO), to manage the information security functions.</li> <li>§ 87% indicated that information security had increased as a business priority at their organizations over the past year.</li> <li>§ Approximately 20 percent of these security incidents ultimately resulted in the loss of patient, financial or operational data.</li> <li>§ Respondents were most likely to be concerned about phishing attacks, negligent insiders and advanced persistent threat (APT) attacks.</li> </ul>
2015 Information Security Breaches Survey	664	<a href="https://www.pwc.co.uk/assets/pdf/2015-isbs-executive-summary-02.pdf">https://www.pwc.co.uk/assets/pdf/2015-isbs-executive-summary-02.pdf</a>	<ul style="list-style-type: none"> <li>§ 90% of large organizations and 74% of small businesses had a security breach (81% and 74% increase respectively from year ago).</li> <li>§ 59% of the respondents expect there will be more security incidents in the next year than last.</li> <li>§ 69% of large organizations and 38% of small businesses were attacked by an unauthorised outsider in the last year (55% and 33% increase respectively from year ago).</li> <li>§ 72% of large organizations and 63% of small businesses provide ongoing security awareness training to their staff (68% and 54% increase respectively from year ago).</li> <li>§ 32% of respondents in 2015 haven't carried out any form of security risk assessment.</li> <li>§ Despite the increase in staff awareness training, people are as likely to cause a breach as viruses and other types of malicious software.</li> </ul>

*continued on following page*

*Table 1. Continued*

Title	Sample Size	Source	Most Important Findings
Cyber security: Are consumer companies up to the challenge? (2014)	111	<a href="https://www.kpmg.com/BE/en/IssuesAndInsights/ArticlesPublications/Documents/Cyber-Security-Survey.pdf">https://www.kpmg.com/BE/en/IssuesAndInsights/ArticlesPublications/Documents/Cyber-Security-Survey.pdf</a>	<p>§ Only 36% of the respondents indicated that their organization has a formal cyber incident response plan.</p> <p>§ Nearly three-quarters of respondents rate their organization's cyber maturity level as average or below.</p> <p>§ The majority of consumer companies are not yet considering how they will respond to a data breach before it occurs.</p> <p>§ 44% of the respondents indicated that in their organization the CIO<sup>1</sup> is responsible for cyber security.</p>
2014 Deloitte-NASCIO Cybersecurity Study	n.a.	<a href="https://www.nascio.org/publications/documents/Deloitte-NASCIOCybersecurityStudy_2014.pdf">https://www.nascio.org/publications/documents/Deloitte-NASCIOCybersecurityStudy_2014.pdf</a>	<p>§ Top five cyber security initiatives for 2014 include: 1/ risk assessments, 2/ training and awareness, 3/ data protection, 4/ continuous security events monitoring, 5/ incident response.</p> <p>§ Lack of sufficient funding continues to be the #1 barrier since 2010.</p> <p>§ Top 3 cyber concerns are as follows: 1/ malicious code (74.5% of respondents), 2/ hactivism (53.2% of respondents), and 3/ zero-day attacks (42.6% of respondents).</p> <p>§ Top 5 cyber threats that the CISOs are more concerned with include: 1/ phishing and pharming; 2/ social engineering; 3/ increasing sophistication of threats; 4/ insecure code; 5/ mobile device threats.</p>
EY's Global Information Security Survey 2014	1825	<a href="https://www.ey.com/Publication/vwLUAssets/EY-global-information-security-survey-2014/\$FILE/EY-global-information-security-survey-2014.pdf">https://www.ey.com/Publication/vwLUAssets/EY-global-information-security-survey-2014/\$FILE/EY-global-information-security-survey-2014.pdf</a>	<p>§ Organizations are lagging behind in establishing foundational cyber security.</p> <p>§ The increased external threat is determined by cyber threats multiplying, disappearing perimeter, and growing attacking power of cyber criminals.</p> <p>§ The increased internal pressures are defined by lack of agility, lack of budget, and lack of skills.</p> <p>§ 56% of organizations say that it is unlikely or highly unlikely that their organization would be able to detect a sophisticated attack.</p> <p>§ 36% of respondents do not have a threat intelligence program.</p> <p>§ 58% of organizations do not have a role or department focused on emerging technologies and their impact on information security.</p>
15 <sup>th</sup> Annual 2010/2011 Computer Crime and Security Survey	351	<a href="https://gatton.uky.edu/FACULTY/PAYNE/ACC324/CSISurvey2010.pdf">https://gatton.uky.edu/FACULTY/PAYNE/ACC324/CSISurvey2010.pdf</a>	<p>§ 21.6% of the respondents indicate that they experienced targeted attacks, moreover, 3% experienced more than 10 targeted attacks.</p> <p>§ The most common types of attacks are (in descending order measured by % of the respondents who answered positively) as follows: malware infection, phishing, laptop/mobile device theft, bots on network, denial of service, password sniffing, financial fraud, exploit of wireless network.</p> <p>§ More than half of losses are not due to malicious insiders.</p>

Source: Author's work

## Experience- and Simulation-based Learning

In terms of education, there is increasing consensus (Berge and Verneil, 2002; OECD, 2014; Biggs and Tang, 2011) that beyond knowledge and skills training, learning process should emphasize on the following: (1) developing a mindset which is global; (2) working through a model of cross-cultural reconciliation; and (3) fostering relational skills. This involves, in the field of cyber security education in the field of economics and business: (1) providing knowledge about cyber threat landscape and the assumptions which underlie intruding business practices and social engineering; (2) concentrating on the context of digital business and business ecosystems; (3) at the individual level, assessing the capabilities to recognize and to avoid cyber attacks. Under these conditions teaching is not merely a way of 'covering the curriculum' or transferring the knowledge directly from the 'expert' to the learner but a way of encouraging innovative thinking, creativity and responsibility for the decisions which are taken.

Research suggests that students must do more than just listen. They must read, write, discuss or be engaged in solving problems. Moreover, students must be engaged in such higher-order thinking tasks as analysis, synthesis, and evaluation, to be actively involved. Thus strategies promoting activities that involve students in doing things and thinking about what they are doing may be called active learning. Performing these activities especially in a team environment forces students to take a responsibility for their decisions.

The distinguishing feature of experience-based learning (or experiential learning) is that the experience of the learner occupies central place in all considerations of teaching and learning. This experience may comprise earlier events in the life of the learner, current life events, or those arising from the learner's participation in activities implemented by teachers and facilitators. A key element of experience-based learning is that learners analyse their experience by reflecting, evaluating and reconstructing it (sometimes individually, sometimes collectively, sometimes both) in order to draw meaning from it in the light of prior experience (Andresen, Boud and Cohen, 2001).

Simulation-based learning is a form of active and experience-based learning (or experiential learning). Its distinguishing feature is that the experience of the learner occupies central place in all considerations of teaching and learning. This experience may comprise earlier events in the life of the learner, current life events, or those arising from the learner's participation in activities implemented by teachers and facilitators. A key element of simulation-based learning is that learners analyse their experience by reflecting, evaluating and reconstructing it (sometimes individually, sometimes collectively, sometimes both) in order to draw meaning from it in the light of prior experience (Andresen, Boud and Cohen, 2001).

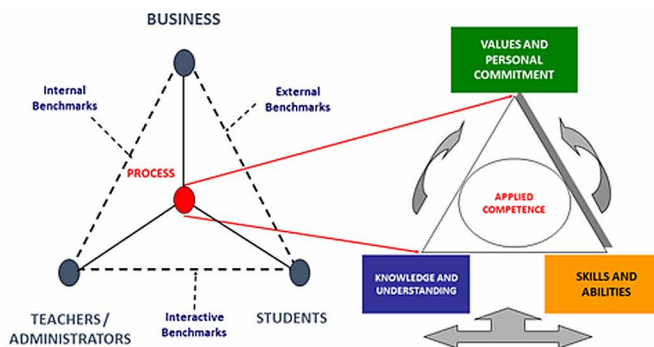
A group of authors (Darling, 1999; Shelton, 1999) proposed that the beginning of the twenty-first century could be called 'The Quantum Age' – time of changing paradigms, from Newton's mechanistic laws of classical physics to the theories of chaos and quantum mechanics. These authors suggest that new sciences provide the conceptual foundation for a new skill set for decision makers – a set of skills that can enable to view conflict from a new perspective, but also to respond to conflict in new ways. This paradigm shift affects the view point to conflicts and respectively to the skills required to deal with conflicts. Shelton and Darling (2004) use quantum theory in their research work as a metaphor for the development of a new set of skills aimed at decision makers, called quantum skills. The concept of quantum skills corresponds to the goals of simulation-based learning and will be used by the authors as a cornerstone of their methodological framework.

## **The Concept of Reference Framework of Applied Competences (REFRAC)**

The concept of REFRAC was built on the assumption of education as a transformative process with its three particular outputs (knowledge, skills and values) which, when linked together, lead to sustainable competence development in any professional setting. None is independent of the others, and it is the interactions among these that leads to sustainability of learning within the profession and competence development. Each domain is, of course, a major field of professional enquiry and action, and its details and form vary from profession to profession.

*Figure 1. REFRAC conceptual model*

*Source: Author's work*



The perspectives on the learning process as a system (teachers and administrators, students, and audience such as policymakers, parents, communities) presented in Figure 1 have their own benefits and standings and also interact with one other continuously. The three major components of this transformative process (presented by the red spot in the middle of the pyramid on Figure 1) are (i) inputs to the educational system, (ii) the system itself, and (iii) the outputs to the system. The inputs to the educational system are the students, faculty and staff, funding, facilities and the university goals. They could be determined as human, physical, and financial resources. The system itself is created and controlled entirely by the elements that compose the system, regardless of the inputs, with some measurable points within; namely, personnel training, teaching methods, learning, advising, counseling, tutoring, evaluations, infrastructure, etc. The system outputs refer to the product that is generated within the system which include tangible outcomes, intangible outcomes and values.



The concept of REFRAC could be easily implemented to cyber security education. It corresponds perfectly to the ISF Security Model (Chaplin and Creasey, 2011) which is developed to support organizations in designing their approach to addressing information security and to give them a basis for identifying the key aspects of an information security programme. Research methodology (Figure 2) includes both qualitative and quantitative methods designed to meet the research goal to identify the critical points of cyber security competences within the background layer of the Reference Framework of Applied Competences (REFRAC).

Figure 2. REFRAC methodology

Source: Author's work

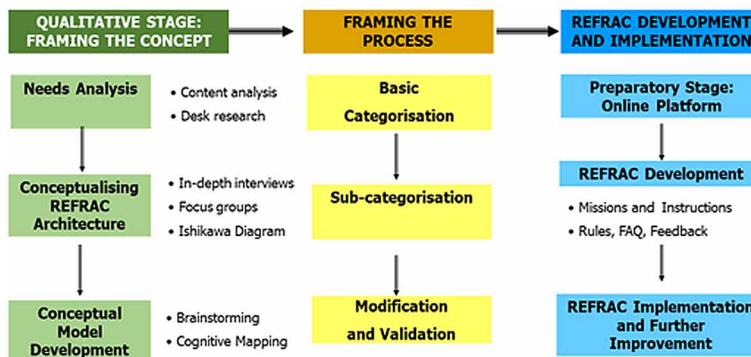
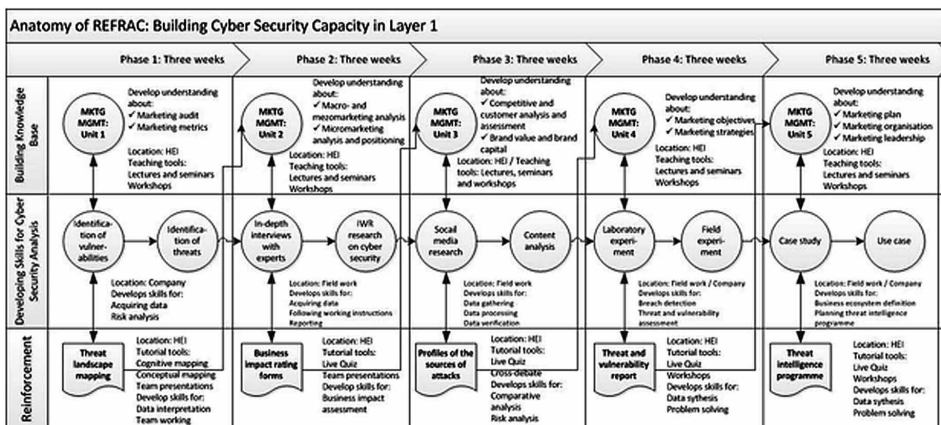


Figure 3. Anatomy of REFRAC

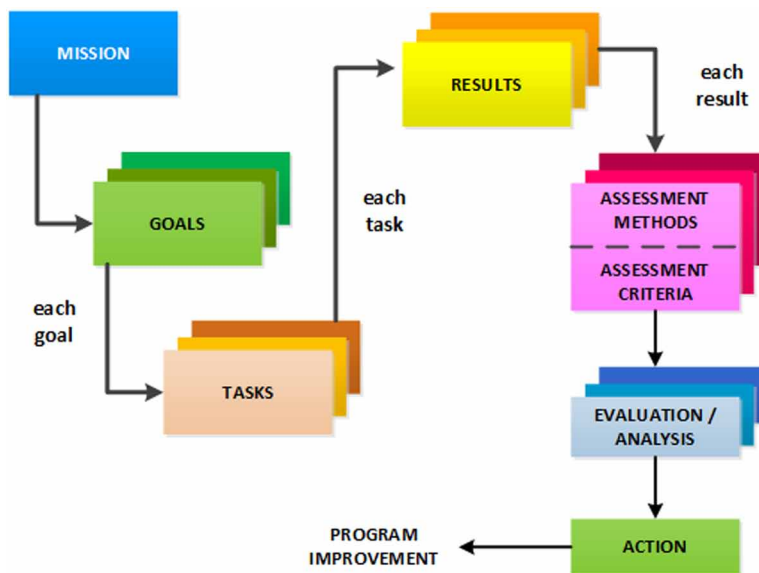
Source: Author's work



The qualitative study involved in-depth interviews with security and cyber security experts, focus group discussions with students, and content analysis to identify the key areas for cyber security capacity building. Five in-depth interviews with experts were conducted followed by two sessions of brainstorming. Two focus groups were undertaken with students. Content analysis was performed in specialised blogs and web posts. As a result a draft conceptual model in a form of Ishikawa diagram was constructed. The model has been “fine-tuned” during a series of workshops. The next step was a content analysis of information published in specialized blogs and web posts. The results showed quite diverse notions toward cyber security education. The most debatable issues covered declarative vs. procedural learning process, technical (computing) vs. non-technical education, traditional vs. interactive, passive vs. active, etc. The author decided that the context of the learning process is critical when the educational goal is to develop applied skills, especially cyber security skills. The students should be willing to explore which requires a combination between contextual and experience-based learning. Quantitative research included annual survey among students to evaluate their awareness and attitudes toward cyber security. The results were used to modify the anatomy of the knowledge areas of REFRAC (Figure 3) and content of the missions.

Mission (Figure 4) is defined as an assignment which requires a practical completion of a task or a sequence of tasks based on a certain knowledge. Its

*Figure 4. Missions: the building blocks of REFRAC*  
*Source: Author's work*

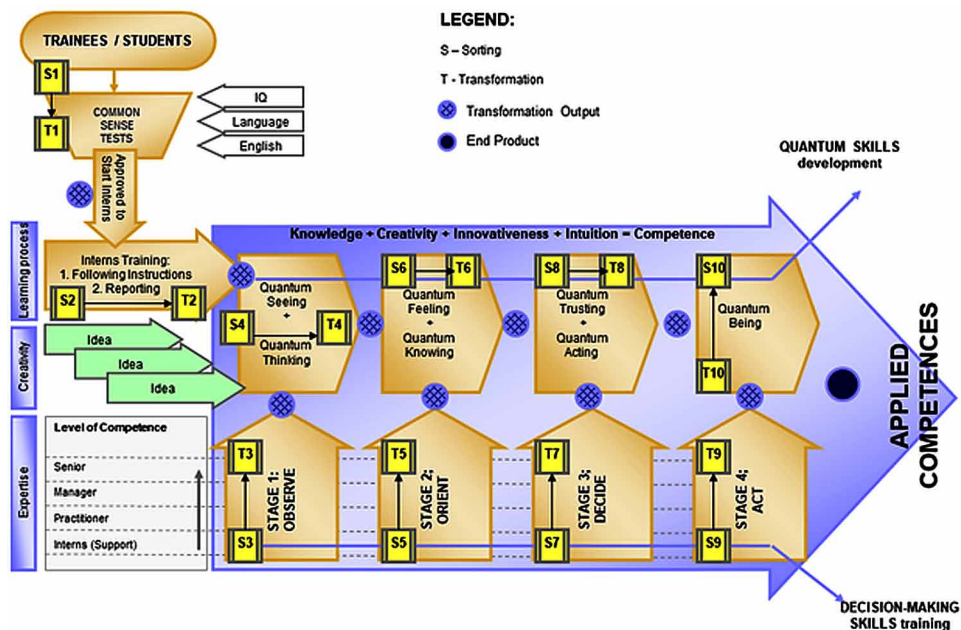


purpose is to crosswalk the skills and knowledge and to stimulate students to make grounded choices. Missions are accompanied by clear instructions and a feedback form. The feedback form is used for validation and it serves as an assessment tool thus providing transparency and creating a competitive environment among students (Vassileva, 2016).

The mission-based methodology flows from the initial mission to the completion of the final mission. REFRAC methodology allows mission re-ordering which depends on planned and expected learning outcomes. REFRAC implementation requires a step-by-step approach. In the beginning of the process students/trainees should pass through initial training on how to read instructions (the core of the missions), how to follow instructions, and how to report the progress and accomplished results. The process itself (Figure 5) combines knowledge development, creativity stimulation, encouragement of innovative thinking, and intuition.

Figure 5. REFRAC simulation-based learning process

Source: Author's work



The focus of REFRAC simulation-based learning process is placed on developing the following three groups of skills. First, cognitive skills which require capability to cope with difficulties provoked by cyber security challenges. These skills are recognized as the main human capabilities, as it requires mental agility and tolerance

for ambiguity or uncertainty to recognize or quickly adapt to the unknown. Second, decision-making skills, especially to understand the true areas of disagreement (conflict) which contribute to solving the right problems and manage the true needs of the parties. Third, tactical abilities. These skills as a background for quantum skills development could be achieved by applying the OODA (Observation – Orientation – Decision – Action) framework (Philip and Martin, 2009). Observation is the means by which one collects/registers information about the state of the external world and corresponds to the key area of structures/systems. Orientation comprises the internal processes by which observations are compared with prior knowledge and experience to update an understanding of the world. It corresponds to the key area of attitudes. Decision is the internal process by which various tentative solutions are assessed and one selected for action. Action is the process by which the internally constructed solution is applied to the world. It corresponds to the key area of behavior.

## **SOLUTIONS AND RECOMMENDATIONS**

Based on results from conducted research and REFRAC methodology (Figure 3), the following structure and organization of course on cyber security is proposed. The first part of the course will include an introduction to cyber security as well as cyber hybrid warfare. In this section terms and definitions will compose the basic knowledge and terminology needed to understand the lectures. The course will in general, provide theory as well as recent publications and research in the field. The second part of the course will include a deep scanning of network security based on tools and examples taken from real scenarios which industrial decision makers in organizations had to face with. The third part of the course will involve a basic training of programming, mainly focuses on structured programming as well as object oriented. The students in this section will learn how to write software code. These pieces of code, are actually planned to be simple programs that can assist to evaluate and monitor malware that penetrates into the industrial's network. The fourth part of the course will be dedicated to security and encryption – description facilities which are common in many industries as well as academic institutions. The fifth part of the course will include a practical sessions on existing tools and models which are aimed to expose the student to cope with cyber attacks scenarios.

## **FUTURE RESEARCH DIRECTIONS**

The proposed framework of applied competences (REFRAC) provide an opportunity for implementation of experience- and simulation-based learning in various scientific

and business areas, incl. cyber security. Its main advantages include: 1/ stimulating creative and innovative (out-of-the-box) thinking, 2/ developing business-related skills at different levels (the basic level comprises business survival skills), 3/ stimulating entrepreneurial attitudes and activities of students. REFRAC could be used as a training tool to build cyber security competences and capacities. With a proper context (knowledge base, see top row in Figure 3), identity (valuable learning outcomes, see bottom row in Figure 3) and structure (flow of missions) REFRAC could contribute to the goals of the Standard of Good Practice for Information Security namely:

- Form basic skills for understanding and implementing the policies, standards and procedures for cyber security;
- Raise information security awareness;
- Form the basis of cyber security assessment;
- Develop specific cyber security arrangements.

When REFRAC is applied in a constant and systematic manner students can gain personal experience through engaging in various business and research activities related to various scientific fields and business areas. The main barriers during the process of REFRAC implementation could be summarised as follows: 1/ Administrative barriers due to the restrictive internal rules of the HEI; 2/ Misunderstanding of the concept both from the management body of the HEI and lecturers /teachers. Such kind of activities require different type of management and high level of engagement of the teaching staff. 3/ Bureaucratic procedures embedded within the educational system which prolong the process of changes and modifications of teaching materials and the process of learning. 4/ Extremely low level of administrative flexibility.

## **CONCLUSION**

The proposed REFRAC conceptual model provides an opportunity for implementation of simulation-based learning in cyber security area but the proposed methodology could be easily adapted to any AI topic, especially responsible AI, ethical considerations toward AI, etc. Facing the challenges of cyber threats, a combined methodology of training and education should be applied in order to: (1) develop cognitive skills; (2) provide situational knowledge; (3) stimulate critical thinking. Under these conditions teaching is not merely a way of “covering the curriculum” or transferring the knowledge directly from the ‘expert’ to the learner but a way of encouraging initiative, creativity and responsibility for the decisions which

are taken. As a result, the competent graduates will possess diverse educational experiences, will be equipped with all required traditional and new skills, including or together with abilities from domain as cultural intelligence, cyber security and public diplomacy. This will require not just to modify our mindset but also to adapt fast to the changing dynamic environment at both individual and institutional level.

## REFERENCES

- Anderson, R. (2008). *Security Engineering: A Guide to Building Dependable Distributed Systems* (2nd ed.). Wiley Publishing, Inc.
- Andresen, L., Boud, D., & Cohen, R. (2001). Experience-Based Learning. In G. Foley (Ed.), *Understanding Adult Education and Training* (2nd ed., pp. 225–239). Allen & Unwin.
- Baryshnikov, Y. (2012). *IT security investment and Gordon-Loeb's 1/e rule*. WEIS paper.
- Berge, Z., & Verneil, M. (2002). The increasing scope of training and development competency. *Benchmarking*, 9(1), 43–61. doi:10.1108/14635770210418579
- Biggs, J., & Tang, C. (2011). *Teaching For Quality Learning At University* (4th ed.). New York: McGraw Hill Society for Research into Higher Education.
- Bissell, K., Lasalle, R., & Dal Cin, P. (2020). *Innovate For Cyber Resilience*. Accenture.com. Available at: [https://www.accenture.com/\\_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf](https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf)
- Buffomante, T. (2020). *All Hands On Deck: Key Cyber Security Considerations For 2020*. KPMG. Available at: <https://home.kpmg/xx/en/home/insights/2020/03/key-cyber-security-considerations-for-2020.html>
- Castelli, C., Gabriel, B., Yates, J., & Booth, P. (2020). *Strengthening Digital Society Against Cyber Shocks*. PwC. Available at: <https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey/strengthening-digital-society-against-cyber-shocks.html>
- Chaplin, M., & Creasey, J. (2011). *The 2011 Standard of Good Practice for Information Security*. Information Security Forum.
- Contu, R., Canales, C., & Pingree, L. (2014). *Forecast: Information Security*. Worldwide, 2012–2018, 2Q14 Update. Gartner report, Gartner, Inc.

Darling, J. R. (1999). Organizational excellence and leadership strategies: Principles followed by top multinational executives. *Leadership and Organization Development Journal*, 20(6), 309–321. doi:10.1108/01437739910292625

Dias, J., Khanna, S., Paquette, C., Rohr, M., Seitz, B., Singla, A., Sood, R., & van Ouwerkerk, J. (2017). *Introducing The Next-Generation Operating Model*. Available at: <https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/introducing%20the%20next-generation%20operating%20model/introducing-the-next-gen-operating-model.ashx>

FireEye. (2020). *Security Effectiveness 2020: Deep Dive Into Cyber Security Reality*. <https://content.fireeye.com/security-effectiveness/rpt-security-effectiveness-2020-deep-dive-into-cyber-reality>

Guzdial, M. (2015). *Using learning sciences to inform cyber security education*. Georgia Tech College of Computing. <https://computinged.wordpress.com/2015/05/18/using-learning-sciences-to-inform-cyber-security-education/>

Harley, D. (2010). *Re-floating the Titanic: dealing with social engineering attacks*. <https://smallbluegreenblog.files.wordpress.com/2010/04/eicar98.pdf>

Herzog, P. (2003). *OSSTMM 2.1 Open-Source Security Testing Methodology Manual*. Institute for Security and Open Methodologies.

Jurse, M., & Mulej, M. (2011). The complexities of business school alignment with the emerging globalisation of business education. *Kybernetes*, 40(9/10), 1440–1458. doi:10.1108/03684921111169477

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The Internet Of Things: Mapping The Value Beyond The Hype*. McKinsey Global Institute. Available at: <https://www.mckinsey.com/business-functions/business-technology/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world>

NIST Roadmap for Improving Critical Infrastructure Cybersecurity. (2014). <https://www.nist.gov/cyberframework/upload/roadmap-021214.pdf>

OECD. (2014). *HEInnovate: Introduction to HEInnovate and its seven dimensions*. Available at: <https://www.oecd.org/cfe/leed/HEInnovate-Introduction%20.pdf>

Philp, W., & Martin, C. (2009). A philosophical approach to time in military knowledge management. *Journal of Knowledge Management*, 13(1), 171–183. doi:10.1108/13673270910931242

### ***Developing Cyber Security Competences Through Simulation-Based Learning***

RSAC. (2019). *The Future of Companies and Cybersecurity Spending*. RSA Conference. Available at: <https://www.rsaconference.com/industry-topics/blog/the-future-of-companies-and-cybersecurity-spending#:~:text=In%202019%2C%20worldwide%20spending%20on,to%20reach%20over%20%24124%20billion.&text=Spending%20on%20security%20services%20has%20reached%20%2464.2%20million%20in%202019>

Shelton, C. (1999). *Quantum Leaps*. Butterworth-Heinemann.

Shelton, C., & Darling, J. (2004). From chaos to order: Exploring new frontiers in conflict management. *Organization Development Journal*, 22(3), 22–41.

Tipton, W.H. (2014). Cyber security education: remove the limits. *Information Week*. <https://www.informationweek.com/government/cybersecurity/cyber-security-education-remove-the-limits/a/d-id/1306950>

Vassileva, B. (2016). Increasing cyber security competences through mission-based learning. *Proceedings of the International Conference on Human Systems Integration Approach to Cyber Security*, 189-204.

World Economic Forum. (2017). *The Global Risks Report 2017 12<sup>th</sup> Edition*. Available at: [http://www3.weforum.org/docs/GRR17\\_Report\\_web.pdf](http://www3.weforum.org/docs/GRR17_Report_web.pdf)

## **ENDNOTE**

<sup>1</sup> CIO – Chief Information Officer.



## Section 3

# Responsible and Ethical Considerations for AI Deployment

## Chapter 9

# The Influence of COVID-19 Outbreak on the Readiness of Firms to Cyber Threats

**Moti Zwillling**  
*Ariel University, Israel*

### ABSTRACT

*Technology impacted the lives of millions of people with their home day-to-day activities. When the COVID-19 pandemic struck in many countries, there was a need to change both the mode of working with technology as well as to handle internet and online risks exposure. During the pandemic, cybercrime groups utilized the internet usage to commit cybercrimes especially by exploiting vulnerabilities of many applications, networks, and infrastructures. This study aims to explore the impact of COVID-19 on the readiness of organizations to handle cyber threats in two directions: 1) analysis of CVE common vulnerability data before and during the pandemic period and 2) analysis of fuzzy logic data model designed to demonstrate the importance of firms readiness to cope with cyber threats. Results show that due to the significant increase in cyber threats, small firms tend to be more fragile to cyber threats than big ones, and they have to invest more resources to mitigate cyber threats. Findings and implications are discussed.*

### INTRODUCTION

Technology has turned to be part of the modern life. Computers, Mobile devices and the Internet Of Things are used for data processing, performing various tasks and for decision making. The internet revolution enabled people around the globe to seek

DOI: 10.4018/978-1-7998-4285-9.ch009

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

for information by using search engines, to meet people in virtual world or through virtual forums and to manage their day life by synchronizing many devices together, to share and distribute information. This phenomena, is also found in companies, which use different systems and technologies to manufacture their products, improve the relationships with their customers through various systems and assist with a faster decisions making. One of the most recent example for the usage of technology by individuals and firms is attributed to the Covid-19 pandemic, in which many people and managers had suddenly forced to use communication tools conducted by the internet for various purposes such as: Work Meetings, Education, Discussions and also for social reasons. The internet which had served as a platform that enables such communication had suddenly become a comfort zone for hackers to commit cyber-crime and hack to sensitive data and systems of individuals and firms. Zwilling et al., (2020, P.1) had already shown that “Internet users possess adequate cyber threat awareness but apply only minimal protective measures usually relatively common and simple one”. In another study, Vassileva and Zwilling (2018), showed that new framework for learning is needed to be part of the modern education among students and employees to better implement and understand how to acquire the knowledge to handle cyber hazards. Lately, Buil-Gil et al., (2020) presented preliminary analysis related to the short term effect of Covid-19 on the cyber dependent crime and online fraud in the United Kingdom. The authors showed a dramatic increase in the number of incidents which were ascribed to cyber-crime and fraud especially in domains of social media, shopping email and auctions. The authors mentioned that incidents were attributed mainly to independent individuals who have low awareness to cyber rather than to employee’s organizations,. This finding is attributed with low willingness of individuals to invest in cyber protection tools, as mentioned by Zwilling et al., (2020). One of the reasons for the increase in cyber-crime incidents due to the Covid-19 pandemic is also related to the transferer from physical environments to online environments in many disciplines such as education, shopping habits, medical service and commerce. One of the main challenges in many organizations especially after the Covid-19 pandemic was to guide the CSO’s (Chief Operations Officer) to clarify the guidelines and establish frameworks that will be used to protect the organization’s assets from cyber-attacks.

## **The Impact of Cyber Hazards on Individuals**

The need to be online and consume internet based services, had been increased tremendously in the recent years. The effect of globalization and the utilization of online services had already been mentioned by de Bruijn and Hanssen in 2017 (de Bruijn and Hanssen, 2017). In their work, the authors mention how globalization development as well as the increase of cloud computing consumption influenced

the way people use technology and change their work habits for private as well as public needs, where in the same time their exposure to cyber incidents is increased as well. Such a phenomena, was shown by Renaud et al., (2020) recently. The authors showed that the Covid-19 pandemic was used to exploit breaches in cyber security in internet application, information systems especially in social media, shopping and cloud services, in which individual's identity and private information was stolen by cyber criminals and groups. This evidence is also mentioned by Gatlan (2020). In his study, the author mentioned that during the period of February-March 2020, a total of 623 million data records were breached by cyber criminals. This finding is also cited by Irwin (2020). When evaluating the cyber victims into groups, one can find three categories: A. Nations or Societies, B. Organizations, C. Individuals (Zhu, Huang and Zhang, 2019). According to Xavir and Pati (2012), Individuals are mainly suffer from phishing attempts, malicious software in addition to an attempt to steal or corrupt their private information. One of the reasons for such occurrence is embodied in the lack of awareness and knowledge which affects individuals readiness to install protection tools to delete malware and/or alert them when cyber incident occurs. Zwillling et al., (2020, p.1) had already emphasized the need for training framework for individuals in order to increase their awareness and knowledge to cyber threats and change their behavior accordingly in order to mitigate the existing risks.

## **The Impact of Cyber Hazards on Organizations**

Organizations and governments are entities that usually deploy wide range of protection tools from cyber-attacks. This may raise the question, if so, what is the reason to find cyber incidents in organizations ?. One of many answers to this question may be that organizations have many computers and information systems that are difficult to maintain during the time. In addition, when more employees use computers and information systems for their work tasks, a constant improvement of systems and training sessions are needed to improve employees awareness to cyber threats as well as the requirement to allocate more budget for this purpose by the firm. Moreover, it was found by many studies, that the weakest chain point with cyber incident is related to the Human Resource, where employees, managers and integrators may implement mistakenly ] policies and instructions which don't follow new frameworks and guidelines to mitigate CVE (Common Vulnerabilities Exploit Solutions). This kind of ignorance may lead into a serious cyber incident which may cost an expense of exceeded money on recovery from cyber incident as well as continuing the business operations. Although, many organizations deliver security training to their employees where their information technology experts provide the assistance with the usage and threats related to cyber, it is found that

especially in the “non-technology departments” (The non-technology back office), still many employees and managers need to refresh the policies and regulations in order to keep their computers and affiliated information “safe” from cyber-attacks.

## **The Implications of Covid-19 on Internet Users**

Cyber Security is a term which holds a wide range of processes and activities to secure individual and organization’s data, networks and applications. Since the Covid-19 pandemic has caused to an enormous increase in the number of cyber-attacks on individuals computer devices as well as on organizations, the immediate effect on firms was reflected in cyber security policies clarification, increase awareness to cyber threats among employees and managers and the need to install update or implement cyber security application tools to mitigate the new posed risks by the Covid-19 pandemic that influenced the mode of work of people around the world. For example: Dowling et al., (2018), had shown that the NIST (National Institute Standards and Technology) framework was used by many firms to mitigate those threats to minimum. As Hakak et al., (2020) shown, the effects of the Covid-19 pandemic on the industry and individuals was resulted in 3 pathways: 1) The effect of the pandemic on the workforce, that was expressed by many people in a way that they had to stay at home and work from distance. The increased demand for internet services consumption such as Cloud Applications and SAAS (Software as a service) and the situation that employees and managers did not have the sufficient knowledge or awareness to cyber threats and how to mitigate them. 2) These threats was found in various sectors that their continues services is dependent on the internet supply ranges from Essential Government Services to Banking Education and Healthcare. Finally 3) The reliance on online platforms had posed cyber security risks such as Financial loss, Corrupted data, Fraud, Data Breach and more.

## **CYBER SECURITY THREATS**

Cyber security threats are divided into many clusters ranges from disrupting services such as DDos attacks, Financial Gains such as Fraud and Ransomware Virus infections, to Data Breach by exploiting applications and system’s vulnerabilities and Firmware attack such as changing the mode of controls, hardware hardens and more. The hacker motivation to break the CIA Triangle (Confidentiality, Integrity and Availability) in order to bypass alert systems and cyber monitoring applications, is driven either by ideological motivation or as in most cases from committing the crime and use the products of the act to gain money. As said cyber security threats are divided into several clusters. I) The DDos attack, is one of the most popular one,

which has been shown by Sharafaldin et al., (2019) as one of the most popular one used by hackers during the Covid-19 pandemic, due to the increase in the number of Internet Users. Hackers, utilized the fact that many people had to work from home, learn by distance and consume online services that are usually conducted in a physical environment (Such as the transfer from Physical Shopping to Online Shopping). The same phenomena was also reported by the Stein and Jacobs (2020), where health services consumption was increased during the first wave of Covid-19 pandemic in the US (around march 2020). The increase in the number of internet users had also influenced the increase in spyware-based programs, as reported by many users around the world. II) The second substantial influence of the Covid-19 pandemic was expressed by an increase of financial gain such as the usage of ransomware viruses (Thakur and Pathan, 2020). In their study, the authors mentioned the ransomware virus as a program that had forced executive managers to pay a ransom to hackers in order to use a special key that will enable them to decrypt content and files which had been encrypted by them and are considered as by a great importance to the business or the perceived reputation of the firm by customers and suppliers. For example: The authors mentioned the “CovidLock” application (Android based) that was used to monitor heat map visuals and statistics during the pandemic. The application was designed in a way that it will be able to grant access to private data installed on the mobile phone. When the grant was provided, the application as said by the authors, locked the pictures, user’s contact, videos and social media content until a ransom money paid in bitcoins (digital currency) by the user. III) Information theft and Data Breach was also observed during the Covid-19 pandemic, where many people around the world such as voice phishing (as called vishing) robocall scams and additional instrument, that as shown by Hobs (2020) were used to steal sensitive information and credentials from users to bypass cyber protection tools. In addition, the most popular phenomena, was the motivation of hackers to seek for vulnerabilities in software tools that are used to perform virtual meetings such as: “Zoom™, Meet Goolge™ etc. and utilize those vulnerabilities to change passwords or steal the user’s credentials. Another malware that is also defined as belonging to the same category is known as “Phishing”. This attack was already reported by Khan and Kahm (2015) as a one that attempt users, especially through e-mails to press or activate a malicious link that activate a malicious software or send a sensitive data to the hacker. During the Covid-19 pandemic, it was reported by Team Risk IQ (2020) that “Over more then 309,000 spam emails that contain either “corona” or “covid” were discovered. IV) Social Networks were also utilized by hackers during the Covid-19 pandemic to deliver disinformation or fake news. As reported by Menon (2020), since the increase in the number of internet usage had been flourished due to the Covid-19 outbreak, social networks such as Facebook™ and LinkedIn™ were

used to deliver disinformation related to the new ways of corona treatment and by delivering a bias news content.

## **METHDOLOGY**

To evaluate the impact of Covid-19 on firms in general and the readiness of firms and individuals in particular to cyber incidents, two studies were conducted: I) Data analysis of CVE (Common Vulnerabilities and Exposures). for the period of years 2000-2020 (Including the months of 3-9/2020) was downloaded and analyzed. II) Simulation of Fuzzy Logic data related to the impact of cyber security breach due to the Covid-19 outbreak on the firm's readiness to handle cyber threats was designed.

### **CVE Data Analysis**

CVE® Files from the Open Source Common Vulnerabilities and Exposures Database was download for analysis. (<https://cve.mitre.org/data/downloads/index.html>).

The file contained the CVE data related to the years of 1999-2020.

The data contained the following attributes:

Date (CVE date), Status (Candidate / Entry) ; CVE Description ; References.

The data was filtered according to specific CVE's types such as – Buffer overflow, SQL injection, and additional common vulnerabilities. The data was filtered with the following years as being included in the study - 2000-2020. Any data for years which are not between this range was excluded. The trend of CVE Exposure and several categories related to Buffer Overflow were analyzed.

The data was analyzed on a macOS system with the Orange™ application for data analytics.

A Fuzzy logic model to evaluate the impact of Covid-19 on the readiness of Firms and to Cyber Incidents was developed by MATLAB R2020a Version.

### **The Fuzzy Logic Model**

A fuzzy logic model based on MATLAB's Fuzzy logic Simulink tool was used. The Advantage of fuzzy logic methodology, is that it enables to represent complex situations and scenarios in which uncertainty is involved. A fuzzy set, as defined by Zadeh (1965) is composed of several elements: Inputs that are connected to outputs through member functions.

The model used is comprised of three steps as described by zwilling (2018):

- “The first step of Fuzzification is executed as the transformation of numerical values into ordinary language, if necessary. For example, the inputs hold linguistic values such as Low, Medium, and High. Each variable usually contains three to seven terms (Attributes).”
- “The second step is expressed in a “Fuzzy Inference” which defines the system behavior by means of the rules such as <if>, <and>, <then>, <with>. The conditional clauses create rules, which evaluate the input variables. These conditional clauses were designed as follows:

*<if>  $x_1$  is value  $x'_1$  <and>  $x'_2$  is value  $x'_2$  ... <and>  $x'_n$  is value  $x'_n$  <then>  $y_1$  is value  $y'_1$  <WITH> probability  $s$ , where  $x_i$  are inputs,  $y_j$  are outputs,  $x'_i$  are values of inputs,  $y'_j$  are values of outputs and  $s$  is a degree of support”.*

- “The third step (“Defuzzification”) is expressed by transformation of linguistic values into numerical ones, if necessary. The outputs also use linguistic values and various types and shapes of membership functions”.

*“The fuzzy logic was described by the following terminology: A fuzzy set  $A$  is defined as  $(U, \mu_A)$ , where  $U$  is the relevant universal set and  $\mu_A: U \rightarrow 0,1$  is a membership function, which assigns each element from  $U$  to the fuzzy set  $A$ . The membership of the element  $x \in U$  of a fuzzy set  $A$  is indicated as  $\mu_A(x)$ . We define  $F(U)$  as the set of all fuzzy set. Then the “classical” set  $A$  is the fuzzy set where:  $\mu_A: U \rightarrow \{0, 1\}$ . Thus  $x \in A$ ,  $\mu_A(x) = 1$  and  $x \notin A$ ,  $\mu_A(x) = 0$ . Let  $U_i$ ,  $i = 1, 2, \dots, n$ , be universals. Then the fuzzy relation  $R$  on  $U = U_1 \times U_2 \times \dots \times U_n$  is a fuzzy set  $R$  on the universal  $U$ ”.*

## **The Design of the Fuzzy Logic Model**

To construct the Fuzzy Interface System by 3 inputs that impact the Cyber Readiness, of a firm during the Covid-19 outbreak, a model design was constructed (Fig 1) with the following Inputs: Firm Type – (Private / Public), Firm Size (Small / Medium / Big), Awareness (Bad / Good / Excellent) and an Output variable: Readiness which has 5 values (Bad / Very Bad / Good / Very Good / Excellent). The rules were set up by the authors. The number of rules was calculated by multiplication of number of terms of inputs (2x3x3x5). We used a trapezoidal shape of membership functions which describes the best way for the searched problem for the input and a triangle shape for the output A symmetrical layout of membership functions was used.

Figures 2 & 3 depicts the model’s parameters and rules graphically and non-graphically



Figure 1. Accumulative CVE Exposures during the years 2000-2020

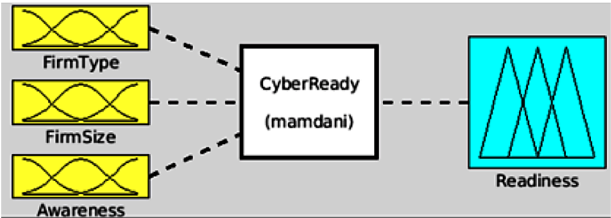
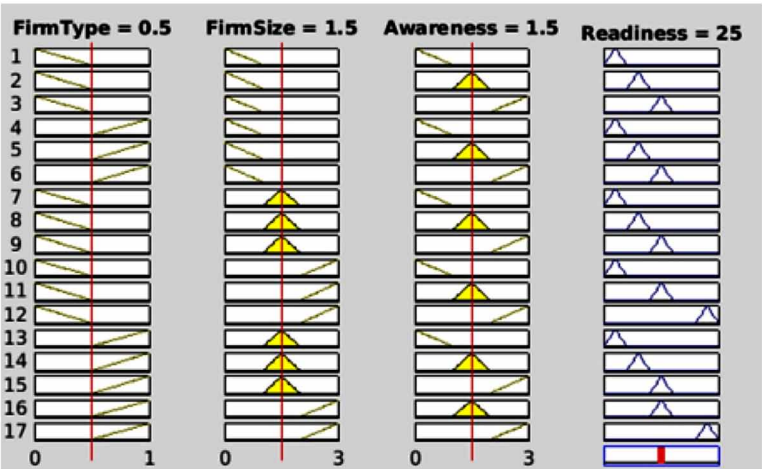


Figure 2. CVE Distribution during the years 2000-2020



The model was tested by using the following numbers: Input Variables: Firm type: Private <0-0.5>, Public <0.5-1> Firm Size: Small <0-1>, Medium <1-2>,>

Figure 3. Frequency of Buffer Overflow Categories along the years 2018-2020 According to 5 categories

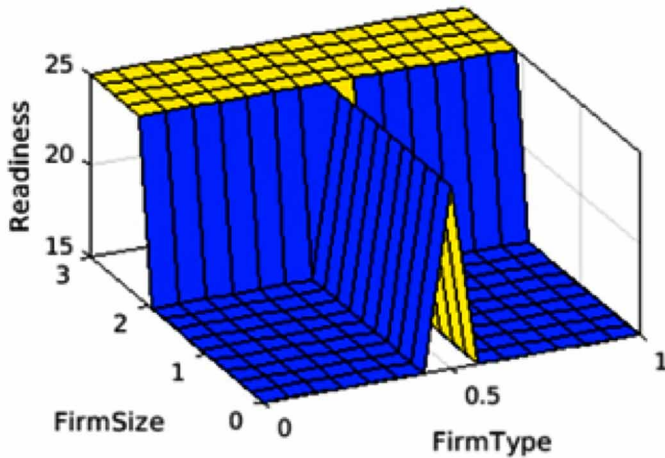
1. (FirmType==Priate) & (FirmSize==Small) & (Awareness==Bad) => (Readiness=Very_Bad) (1)
2. (FirmType==Priate) & (FirmSize==Small) & (Awareness==Good) => (Readiness=Bad) (1)
3. (FirmType==Priate) & (FirmSize==Small) & (Awareness==Excellen) => (Readiness=Good) (1)
4. (FirmType==Public) & (FirmSize==Small) & (Awareness==Bad) => (Readiness=Very_Bad) (1)
5. (FirmType==Public) & (FirmSize==Small) & (Awareness==Good) => (Readiness=Bad) (1)
6. (FirmType==Public) & (FirmSize==Small) & (Awareness==Excellen) => (Readiness=Good) (1)
7. (FirmType==Priate) & (FirmSize==Medium) & (Awareness==Bad) => (Readiness=Very_Bad) (1)
8. (FirmType==Priate) & (FirmSize==Medium) & (Awareness==Good) => (Readiness=Bad) (1)
9. (FirmType==Priate) & (FirmSize==Medium) & (Awareness==Excellen) => (Readiness=Good) (1)
10. (FirmType==Priate) & (FirmSize==Big) & (Awareness==Bad) => (Readiness=Very_Bad) (1)
11. (FirmType==Priate) & (FirmSize==Big) & (Awareness==Good) => (Readiness=Good) (1)
12. (FirmType==Priate) & (FirmSize==Big) & (Awareness==Excellen) => (Readiness=Excellent) (1)
13. (FirmType==Public) & (FirmSize==Medium) & (Awareness==Bad) => (Readiness=Very_Bad) (1)

Big <2-3>; Awareness: Bad <0-1>, Good <1-2>, Excellent <2-3> ; Output: Readiness: Very Bad <0-10>, Bad <10-20>, Good <20-30>, Very Good <30-40>, Excellent <40-50>.

A combination of rules were constructed with the whole combinations: For example:

**Rule I:** “If (Firm Type is Private) and (Firm Size is Small) and (Awareness is Bad) Then (Readiness is Very Bad) (Refer to Figures 2 & 3). A surface view related to variables combinations was constructed (Fig 4).

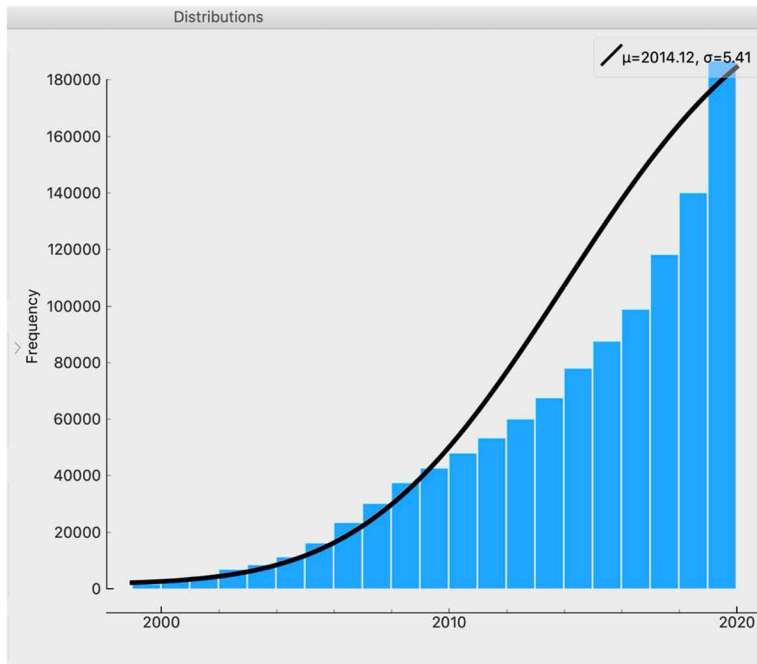
Figure 4. Accumulative Frequency of Buffer Overflow incidents during the period 2018-2020



## STUDY I AND II RESULTS

Figure 5 and Figure 6 exhibit a substantial increase in CVE exposure and accumulative exposure during the years 2000-2020. One of the prominent finding is that since the Covid-19 outbreak it is seen that the frequency of CVE incidents are almost doubled from 20,000 incidents to 40,000. Figure 7, exhibits the frequencies of Buffer overflow from 2018-2020 for 5 categories, where it is seen, that buffer overflow “Can occur”, meaning that less attention was given to buffer overflow issues regarding other cases. Moreover, Figure 8, exhibits the accumulative frequency where it is deduced that, when various buffer overglow categories occur during the year 2020,

*Figure 5. Fuzzy Logic model related to Readiness of a firm as an input of firm type, firm size and readiness to cyber management*



more incidents had emerged then in previous years (especially during the covid-19 First Wave outbreak).

## **DISCUSSION, CONCLUSION AND FUTURE RESEARCH**

Based on recent literature studies it was shown that CVE exposure had increased dramatically during the Covid-19 pandemic. When many firms had to change their mode of work to online, exposure to more risks had occurred which influenced dramatically on their readiness for cyber mitigation then in “Normal” timeline. The CVE Analysis shows that the exposure is reflected by various risks ranging from SQL Injection, Vulnerabilities in applications to Buffer Overflow. In addition, the MATLAB fuzzy model shows that the firm strength is dependent especially on the firm size and the readiness of its employees. Small firms that are private suffer more than big firms. Yet, while this finding is not much surprising the model can assist to mitigate the various cyber risks in a way that it will serve firms to simulate

Figure 6. Fuzzy Logic Rules (Graphically)

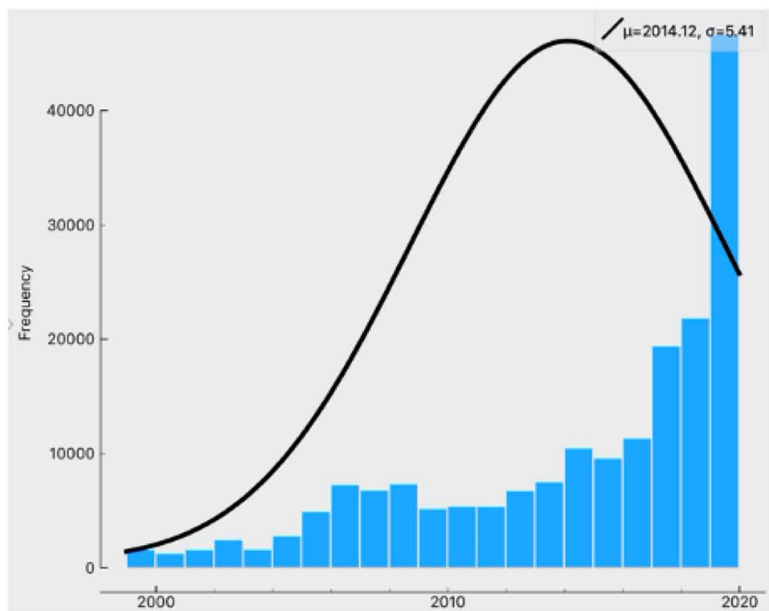
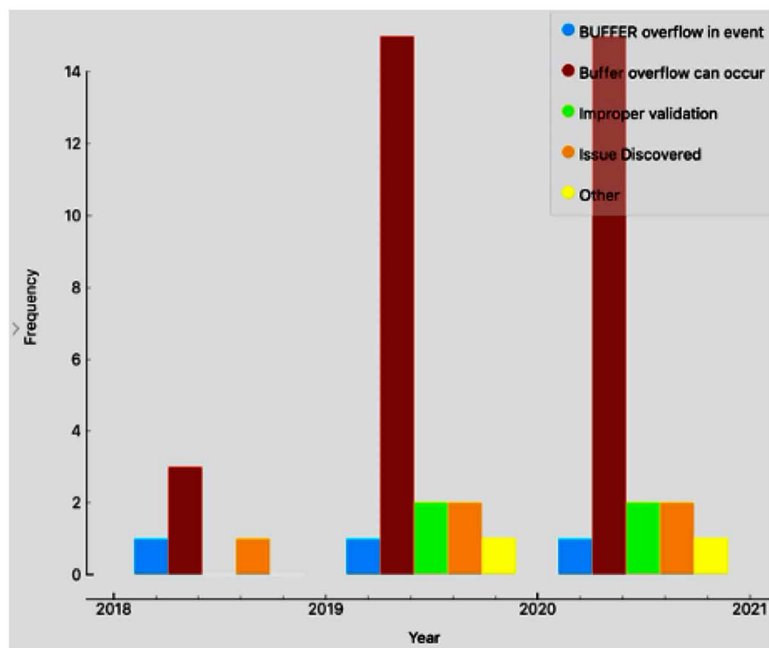
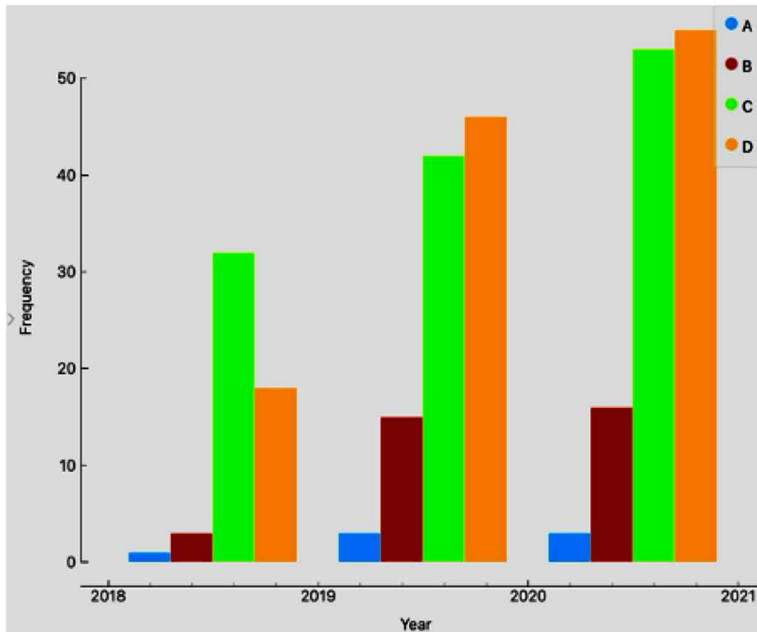


Figure 7. Fuzzy Logic Rules



*Figure 8. Fuzzy Logic Surface*



different scenarios of readiness and identify lack of awareness to cyber processes, applications and procedures among manager and employees.

From all of the above, it can be seen that organizations need to have a systematic vision of cyber readiness. The fuzzy logic and analysis of CVE breaches via applications, networks and systems can assist managers and firm to reduce the cyber risks and exposures especially during difficult mode of work such as the one enforced by the Covid-19 on many different types of organizations.

## REFERENCES

- Buil-Gil, D., Miró-Llinares, F., Moneva, A., Kemp, S., & Díaz-Castaño, N. (2020). Cybercrime and shifts in opportunities during COVID-19: A preliminary analysis in the UK. *European Societies*, 1–13. doi:10.1080/14616696.2020.1804973
- de Bruijn, H., & Janssen, M. (2017). Building cybersecurity awareness: The need for evidence-based framing strategies. *Government Information Quarterly*, 34(1), 1–7. doi:10.1016/j.giq.2017.02.007

- Dowling, S., Schukat, M., & Barrett, E. (2018). Improving adaptive honeypot functionality with efficient reinforcement learning parameters for automated malware. *Journal of Cyber Security Technology*, 2(2), 75–91. doi:10.1080/23742917.2018.1495375
- Gatlan, S. (2020). Coronavirus Phishing Attacks Are Actively Targeting the US. *Bleeping Computer*. <https://www.bleepingcomputer.com/news/security/coronavirus-phishing-attacks-are-actively-targeting-the-us/>
- Hakak, S., Khan, W. Z., Imran, M., Choo, K. K. R., & Shoaib, M. (2020). Have You Been a Victim of COVID-19-Related Cyber Incidents? Survey, Taxonomy, and Mitigation Strategies. *IEEE Access: Practical Innovations, Open Solutions*, 8, 124134–124144. doi:10.1109/ACCESS.2020.3006172
- Hobbs, K. (2020). Socially distancing from COVID-19 robocall scams. *Consumer Information*. <https://www.consumer.ftc.gov/blog/2020/03/socially-distancing-covid-19-robocall-scams>
- Irwin, L. (2020). *List of Data Breaches and Cyber Attacks in February 2020–623 million Records Breached*. IT Governance.
- Khan, W. Z., Khan, M. K., Muhaya, F. T. B., Aalsalem, M. Y., & Chao, H. C. (2015). A comprehensive study of email spam botnet detection. *IEEE Communications Surveys and Tutorials*, 17(4), 2271–2295. doi:10.1109/COMST.2015.2459015
- Menon, S. (2020, April 19). Coronavirus: Herbal remedies in India and other claims fact-checked. *BBC News*. <https://www.bbc.com/news/world-asia-india-51910099>
- Renaud, K., & Weir, G. R. (2016). Cybersecurity and the Unbearability of uncertainty. *2016 Cybersecurity and Cyberforensics Conference (CCC)*. 10.1109/CCC.2016.29
- Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2019). An evaluation framework for network security visualizations. *Computers & Security*, 84, 70–92. doi:10.1016/j.cose.2019.03.005
- Stein, S., & Jacobs, J. (2020, March 16). *Cyber-attack hits U.S. health agency amid COVID-19 outbreak*. Bloomberg.com. <https://www.bloomberg.com/news/articles/2020-03-16/u-s-health-agency-suffers-cyber-attack-during-covid-19-response>
- Thakur, K., & Pathan, A. S. K. (1995, August 10). *Cybersecurity fundamentals: A real-world perspective*. Routledge & CRC Press. <https://www.routledge.com/Cybersecurity-Fundamentals-A-Real-World-Perspective/Thakur-Pathan/p/book/9780367472504>

Vassileva, B., & Zwilling, M. (2018). Hybrid Warfare Simulation-Based Learning: Challenges and Opportunities. *Information & Security*, 39(1), 220–234.

Xavier, U. H. R., & Pati, B. P. (2012, November). Study of internet security threats among home users. In *2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)* (pp. 217-221). IEEE. 10.1109/CASoN.2012.6412405

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. doi:10.1016/S0019-9958(65)90241-X


Zhu, J., Huang, H., & Zhang, D. (2019). Big Tigers, Big Data: Learning Social Reactions to China's Anticorruption Campaign through Online Feedback. *Public Administration Review*, 79(4), 500–513. doi:10.1111/puar.12866

Zwilling, M., Levy, S., Gvili, Y., & Dostal, P. (2020). Machine learning as an effective paradigm for persuasive message design. *Quality & Quantity*, 54(3), 1–23. doi:10.1007/11135-020-00972-0

# Chapter 10

## Ethical Aspects of Information Literacy in Artificial Intelligence


**Selma Leticia Capinzaiki Ottonicar**

 <https://orcid.org/0000-0001-6330-3904>  
*Sao Paulo State University (UNESP), Brazil*

**Ilídio Lobato Ernesto Manhique**

*Sao Paulo State University (UNESP), Brazil*

**Elaine Mosconi**

 <https://orcid.org/0000-0001-5579-9997>  
*Université de Sherbrooke (UdeS), Canada*

### ABSTRACT

*The purpose is to analyse information literacy to provide ethical insight into artificial intelligence. The methodology was based on a systematic literature review of SCOPUS, Web of Science, Library and Information Science Abstracts, and Science Direct. The results demonstrated that there are only a few studies about the topic, so there is a research opportunity about this type of literacy and its ethical aspects in the context of artificial intelligence. As a conclusion, information literacy is crucial to the development of critical thinking in technology use. Information literacy should be applied in artificial intelligence courses to discuss ethical aspects of technology.*

DOI: 10.4018/978-1-7998-4285-9.ch010

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



## **INTRODUCTION**

The fourth industrial revolution or Industry 4.0 (I4.0) is based on the connection between machines, networks and humans (Adolphs & Epple, 2017; Almada-Lobo, 2015; Schwab, 2016). One part of I4.0 is artificial intelligence (McCarthy et al., 1956), which has improved individuals' lives through the human-machine interaction. Despite its potential, artificial intelligence has created some ethical concerns. This technology has implications for cybersecurity and data privacy. Therefore, there is a dichotomy because artificial intelligence influences society both positively and negatively.

Artificial intelligence is a topic of research of many fields, so it can be considered multidisciplinary. The topic of this chapter is artificial intelligence in the context of Ethics, Information Science and Computer Science fields. Only a few researches have studied artificial intelligence in a multidisciplinary perspective, hence this chapter helps to fill this knowledge gap. Information Science is a useful lens, since it values the union of different knowledge.

Information Literacy is a very well-known topic in the field of Information Science. It is the ability to access, evaluate and use information to construct critical thinking. Ethics is a relevant aspect of information literacy (Vitorino & Piantola, 2011) because it guides individuals' behavior in society. An ethical approach to information has been recognized by Information Literacy Competency Standards for Higher Education of American College & Research Association (ACRL). This organization considers that an information literate person understands economic, legal and social aspects of information use. Furthermore, ACRL (2000) understands that information literacy helps individuals to access and use information in an ethical and legal way.

Information literacy allows people to interpret information issues in a critical way (Belluzzo, 2014; Yafushi, 2015; Ottonicar, Valentim & Feres, 2016) and it is the means through which people experience information (Demasson, Patritdge & Bruce, 2016). It helps to develop critical thinking (Grafstein, 2017), so it is a sociocultural element to allow individuals to deal with complex contexts (Lloyd, 2007).

Many countries have introduced laws and public policy focused on information literacy. However, those actions are still limited, because the obscurity of the current social system can create illusions in peoples' minds (Slayton, 2018). The use of ethical aspects of information literacy (Vitorino & Piantola, 2011) is the first step to face the dichotomy of artificial intelligence. Professionals need to understand the consequences of artificial intelligence to society. Bostron (2016) criticizes our society and compares humans to "children who play with a bomb".

Based on these ideas, this book chapter has the following question: how can information literacy contribute to an ethical development of artificial intelligence?

The purpose of this chapter is to analyse information literacy to provide ethical insight into artificial intelligence. Furthermore, this study aims to discuss the contribution of information literacy to the ethical use of artificial intelligence and to identify authors who have researched this topic.

The chapter is structured in six sections. First, it presents the introduction, research questions and purpose of the chapter. Second, it shows the method, systematic literature review, and some quantitative results. Third, the chapter discusses artificial intelligence and its ethical issues. Fourth, a section about information literacy concepts and ethics is introduced in the literature review. Fifth, the results showed the papers retrieved and ideas of authors. In the final section, this chapter explains the conclusions, limitations and future research.

## **BACKGROUND**

### **Artificial Intelligence and Ethical Issues**

In today's society, there are some utopic ideas about the benefits of technology. In a perfect world, information and communication technology would be the source of transformation to society. That optimism is shared by some authors such as Bell (1973). He believes that technology has an important role to determine people's lives.

According to Lyon (1992) the relationship between technology and society cannot be hierarchical and vertical, which means that technology and individuals should be seen as equally important. The author (1992) explains that technology and society have an interdependent relationship because technology influences social relationships and it is also a social product.

The development of technology leads to some ethical problems related to the use of knowledge. Morin (2005) provides an example of these dangers: nuclear production contributed to national socioeconomic development, but it also nearly caused the annihilation of humanity in the mid-20<sup>th</sup> century. People need to abandon the naive idea of good or bad science and technology. The dichotomy of technology must be discussed from an ethical and moral perspective in order to understand how information literacy can contribute to artificial intelligence development.

Artificial intelligence emerges from technological changes in contemporary society. This intelligence has challenged traditional ideas of time, space and intelligence. These ideas can also be discussed by Philosophy of Information, which is a new field of investigation of Information Science and Computer Science (Floridi, 2004). Philosophy of Information involves the critical nature of concepts and basic principles of information. Furthermore, it helps to develop information theory and computer methodology applied to philosophical problems (Floridi, 2004).

Floridi (2004) discusses artificial intelligence in the information processing view, and he considers that artificial intelligence studies human cognition and information systems. Information processing, cognition and intelligence are the most relevant problems of Philosophy of Information. One of the problems is: Are there types of cognition? Can cognition be understood in a non-biological perspective? There is a relational perspective between human and computer intelligence which is a more abstract and holistic way of building symbols upon reality. This idea presupposes a distinction between natural and artificial intelligence (Floridi, 2004).

According to that author, artificial intelligence is limited to data processing (non-interpreted patterns of differences and deviations) while natural intelligence mainly processes information (patterns that are well-formed out of significant data). So artificial intelligence should be described as a data system and natural intelligence as an information system.

Araújo (2017) discusses the consequences of artificial intelligence to produce both information and knowledge, since algorithms have been used to create information. This author (2017) analyses artificial intelligence in many aspects of information production. He also studies the challenges of the knowledge production field in the context of artificial intelligence.

In the United States a business called Automated Insights uses a software program called Wordsmith to create journalistic texts. Some newspapers have published these articles, naming Automated Insights as the author. These algorithms are a topic for ethical discussion that includes the fear of robots replacing journalists (Araújo, 2017). Philip Parker is the author of dictionaries, financial reports, didactic books, and medical texts that are created by an algorithm. His books are on sale at Amazon. He constructs the steps to write a technical text that imitates an academic researcher (Araújo, 2017).

The impacts of information technology have caused some ethical issues. Because of that, many organizations have created a set of solutions. These solutions aim to guide professionals, such as with rules and laws that regulate the use of artificial intelligence. The generation of texts by technologies worry academics, because technology can influence the way people write papers in the future (Araújo, 2017, p. 90). The development of algorithms contributes to new demands in society. First, algorithms demand the possibility of artificial intelligence of transforming science and second, algorithms demand individuals to become lifelong learners. Furthermore, this context requires the development of abilities to evaluate information sources, so people can identify the author of a text.

Information literacy is lifelong learning through the critical interpretation of information. This literacy is also focused on the ethical use and dissemination of information. The increase of artificial intelligence tools will cause many challenges to scientists, especially concerning the authorship and origin of scientific papers

(Araújo, 2017). According to Araújo (2017, p.95): “In the future, plagiarism will become a smaller problem. The greatest challenge will be to know if researchers are truly the authors of the papers they submit or if they should be considered only meta-authors of their respective studies. The question is: Who is the author of a paper? The algorithm that created the text or the person who programmed the algorithm? The answer to this question implies a re-evaluation of the concept of authorship of academic papers”.

Algorithms are created using a system called deep learning that allows automatic correction without the programmer’s intervention (Araújo, 2017). Since they correct themselves, could they become authors of scientific papers? According to Capurro (2007) and Hjørland (2004) knowledge depends on the interaction between cognitive and social interaction. Technology can generate data while humans create knowledge. In this book chapter, we emphasize that humans can produce knowledge based on the interactions with other people and technology. Technology alone cannot yet construct knowledge, since knowledge depends on social, cultural and moral aspects.

## **Information Literacy**

Information literacy is a set of abilities to evaluate information sources. Information can be manipulated, so people need to be information literate to interpret that information and identify fake news. In contemporary society, individuals have access to a lot of information, and they need to know how to access, filter, organize and share information.

Information literacy was coined in 1974 by Paul Zurkowski. Since then, researchers have studied the topic in many fields such as workplaces, libraries, schools, industry, community, ethics and art. Therefore, Information literacy is interdisciplinary because its concept is created by different fields. In 2009 the then-President of United States Barack Obama announced Information Literacy Day. Obama helped to disseminate this literacy worldwide to become more valuable to society. The document says:

*Though we may know how to find the information we need, we must also know how to evaluate it. Over the past decade, we have seen a crisis of authenticity emerge. We now live in a world where anyone can publish an opinion or perspective, whether true or not, and have that opinion amplified within the information marketplace. At the same time, Americans have unprecedented access to the diverse and independent sources of information, as well as institutions such as libraries and universities, that can help separate truth from fiction and signal from noise (Obama, 2009).*

One of the fields of research is ethics. Information literacy helps people to become more ethical when using information. Information should be used to improve human

knowledge and socialization. Information literate people respect the rights of others and value ethics in an artificial intelligence context.

The set of moral values of a group can be evaluated and interpreted in a critical way. Information literacy contributes to this issue, helping people to learn how to behave according to moral values and become an example to others. Individuals can judge right from wrong in many aspects of their lives.

Furthermore, individuals can understand the consequences of information use to society, and they need to know how to structure information (Bembem & Santos, 2014).

The document Framework of Information Literacy for Higher Education, released in 2000, updated the information literacy standards (ACRL, 2015). Furthermore, the ACRL (2015) also changed the concepts of information literacy, because they introduced individuals' experience as an element of learning. The ACRL (2015) document is called Information Literacy Competency Standards for Higher Education which demonstrates new concepts with a prescriptive interpretation.

Information literacy is the set of integrated abilities encompassing the reflective discovery of information, the understanding of how information is produced and valued, and the use of information in creating new knowledge and participating ethically in communities of learning (ACRL, 2015, p. 11).

The Chartered Institute of Library and Information Professionals (CILIP) in the United Kingdom also contributed to the development of the information literacy concept. Initially, information literacy was considered the ability to identify information needs, access it, evaluate it, use it, and communicate it ethically. In 2018 the CILIP expanded the concept of information literacy to consider different contexts of learning: *"Information literacy is the ability to think critically and make balanced judgements about any information we find and use. It empowers us as citizens to develop informed views and to engage fully with society"* (CILIP, 2019).

Information literacy is a process that helps people to face ethical, legal and political problems (Marti & Vega-Almeida, 2005). In order to solve these problems, information literacy has been inserted into school curricula. Some organizations have disseminated information literacy standards and indicators, such as the American Association of School Libraries (1998) and the International Federation of Libraries Association (IFLA). The American Association of School Libraries (1998) suggests the following standards:

Standard 1 The student who is information literate accesses information efficiently and effectively.

Indicator 1. Recognizes the need for information

Indicator 2. Recognizes that accurate and comprehensive information is the basis for intelligent decision making

## ***Ethical Aspects of Information Literacy in Artificial Intelligence***

Indicator 3. Formulates questions based on information needs

Indicator 4. Identifies a variety of potential sources of information

Indicator 5. Develops and uses successful strategies for locating information

Standard 2 The student who is information literate evaluates information critically and competently.

Indicator 1. Determines accuracy, relevance, and comprehensiveness

Indicator 2. Distinguishes among fact, point of view, and opinion

Indicator 3. Identifies inaccurate and misleading information

Indicator 4. Selects information appropriate to the problem or question at hand

Standard 3 The student who is information literate uses information accurately and creatively.

Indicator 1. Organizes information for practical application

Indicator 2. Integrates new information into one's own knowledge

Indicator 3. Applies information in critical thinking and problem solving

Indicator 4. Produces and communicates information and ideas in appropriate formats

Standard 4 The student who is an independent learner is information literate and pursues information related to personal interests.

Indicator 1. Seeks information related to various dimensions of personal well-being, such as career interests, community involvement, health matters, and recreational pursuits

Indicator 2. Designs, develops, and evaluates information products and solutions related to personal interests

Standard 5 The student who is an independent learner is information literate and appreciates literature and other creative expressions of information.

Indicator 1. Is a competent and self-motivated reader

Indicator 2. Derives meaning from information presented creatively in a variety of formats

Indicator 3. Develops creative products in a variety of formats

Standard 6 The student who is an independent learner is information literate and strives for excellence in information seeking and knowledge generation.

Indicator 1. Assesses the quality of the process and products of personal information seeking

Indicator 2. Devises strategies for revising, improving, and updating self-generated knowledge

Standard 7 The student who contributes positively to the learning community and to society is information literate and recognizes the importance of information to a democratic society.

Indicator 1. Seeks information from diverse sources, contexts, disciplines, and cultures

Indicator 2. Respects the principle of equitable access to information

Standard 8 The student who contributes positively to the learning community and to society is information literate and practices ethical behavior in regard to information and information technology.

Indicator 1. Respects the principles of intellectual freedom

Indicator 2. Respects intellectual property rights

Indicator 3. Uses information technology responsibly

Standard 9 The student who contributes positively to the learning community and to society is information literate and participates effectively in groups to pursue and generate information.

Indicator 1. Shares knowledge and information with others

Indicator 2. Respects others' ideas and backgrounds and acknowledges their contributions

Indicator 3. Collaborates with others, both in person and through technologies, to identify information problems and to seek their solutions

Indicator 4. Collaborates with others, both in person and through technologies, to design, develop, and evaluate information products and solutions.

Association of College & Research Libraries (ACRL, 2015):

Standard 1 - The information literate student determines the nature and extent of the information needed.

Standard 2 - The information literate student accesses needed information effectively and efficiently.

Standard 3 - The information literate student evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.

Standard 4 - The information literate student, individually or as a member of a group, uses information effectively to accomplish a specific purpose.

Standard 5 - The information literate student understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally.

Council of Australian University Librarians (CAUL, 2004):

Standard 1 - The information literate person recognises the need for information and determines the nature and extent of the information needed

Standard 2 - The information literate person finds needed information effectively and efficiently

Standard 3 - The information literate person critically evaluates information and the information seeking process

Standard 4 - The information literate person manages information collected or generated

Standard 5 - The information literate person applies prior and new information to construct new concepts or create new understandings

Standard 5 - The information literate person uses information with understanding and acknowledges cultural, ethical, economic, legal, and social issues surrounding the use of information

IFLA (Lau, 2008, p.16-17) information literacy standards are as follows:

1. **Access.** The user accesses information effectively and efficiently
  - a. Definition and articulation of the information need
    - i. Defines or recognizes the need for information
    - ii. Decides to do something to find the information
    - iii. Express and defines the information need Initiates the search process
  - b. Location of information
    - i. Identifies and evaluates potential sources of information
    - ii. Develops search strategies Accesses the selected information sources  
Selects and retrieves the located information
2. **Evaluation.** The user evaluates information critically and competently
  - a. Assessment of information
    - i. Analyzes, examines, and extracts information
    - ii. Generalizes and interprets information
    - iii. Selects and synthesizes information
    - iv. Evaluates accuracy and relevance of the retrieved information
  - b. Organization of information Arranges and categorizes information
    - i. Groups and organizes the retrieved information
    - ii. Determines which is the best and most useful information
3. **Use.** The user applies/uses information accurately and creatively
  - a. Use of information
    - i. Finds new ways to communicate, present and use information
    - ii. Applies the retrieved information
    - iii. Learns or internalizes information as personal knowledge
    - iv. Presents the information product
  - b. Communication and ethical use of information
    - i. Understands ethical use of information
    - ii. Respects the legal use of information
    - iii. Communicates the learning product with acknowledgement of intellectual property
    - iv. Uses the relevant acknowledgement style standards



## MAIN FOCUS OF THE CHAPTER

### Method

This is an exploratory and descriptive research, based on a Systematic Literature Review (SLR). The SLR aims to identify, select, evaluate, synthesize and reproduce relevant evidence about specific objects (Tranfield, Denyer & Smart, 2003). This method follows some steps that validate the results as illustrated by Table 1.

*Table 1. Phases of SLR*

Step	Phase
1. Planning of SLR	Identify the need for revision Prepare the proposal Develop the protocol of SLR
2. Application of SLR	Identify studies Select research Evaluation of the quality of the study Data synthesis Data extraction
3. Sharing the results	Recommendations and results Demonstrate evidence that is useful to practical contexts.

Source: Adapted from Tranfield, Denier and Smart (2003) and Bordeleau, Mosconi and Santa-Eulalia (2018, p. 3946).

The first step of the research uses a bibliographic review to construct the concepts of information literacy and artificial intelligence focused on ethics. The second step was to access international scientific databases such as *Base de Dados de Artigos de Periódicos da Ciência da Informação Brasileira* (BRAPCI) [Brazilian Information Science Database of Journal Papers] which is an important database of Information Science. The database search only found one paper that connected information literacy to artificial intelligence.

SCOPUS, Web of Science (WoS), Library and Information Science Abstracts (LISA) and Science Direct are also part of the SLR in this chapter. These databases are very important to science and they connect many fields of research. The data gathering was based on advanced research in each database because that helps to find precise results. The SLR was implemented using the expressions, “information literacy AND artificial intelligence”. The papers were retrieved from any year, that is, there were no time constraints.

The paper selection followed these criteria: (i) academic papers or scholarly papers; (ii) papers that are peer reviewed; and (iii) mention of both artificial intelligence and information literacy together in one of the following sections of the retrieved article: title, abstract, and keywords. After the analysis of papers (SCOPUS, WoS, LISA and Science Direct), the research generated these results:

*Table 2. Quantitative results of SLR*

Database	Papers Available
SCOPUS (Elsevier)	0
WoS (Clarivate Analytics)	10
Science Direct (Elsevier)	5
LISA	6
BRAPCI	1

Source: The authors - 2019

The papers retrieved from the databases were saved in RIS format, so they could be organized at Endnote Software. This software helped to analyze papers based on the inclusion criteria. After removing duplicates and filtering by topic relevance, this research analyzed only 15 papers. Therefore, the SLR was based on fifteen scientific papers. This demonstrates that there is a lack of research connecting information literacy to the context of artificial intelligence.

**Results**

**Systematic Literature Review**

The results of the SLR show that there is not a lot of research connecting information literacy and artificial intelligence. This type of research is still innovative to Information Science and to other interdisciplinary fields such as Computer Science, Education, etc.

This finding is supported by the fact that SCOPUS, an important scientific database, did not find any paper about the topic. The other databases (Web of Science, Science Direct and LISA) retrieved 14 (fourteen) papers that connect information literacy to artificial intelligence (Table 3 below).

The papers in Table 3 are connected to four scientific fields: Information Science, Computer Science, Communication and Education. These fields are multi- and

Table 3. Papers retrieved in databases

Authors	Title	Periódico/ Tipo de Documento	Year
Mathews, E. C.; Jackson, G. T.; Olney, A.; Chipman, P.; Graesser, A. C.	Achieving domain independence in AutoTutor	Book Chapter	2003
Mathews, E. C.; Jackson, G. T.; Olney, A.; Chipman, P.; Graesser, A. C.	Achieving domain independence in AutoTutor	Book Chapter	2003
Terracina, A.; Mecella, M	Building an Emotional IPA Through Empirical Design With High-School Students	Book Chapter	2015
Forbes, J.; Garcia, D. D.	But What Do the Top-Rated Schools Do?" A Survey of Introductory Computer Science Curricula	Book Chapter	2007
Raycheva, L.	The digital notion of the citizen-centred media ecosystem	International Journal of Digital Television	2018
Hesse, B. W.; Shneiderman, B.	eHealth research from the user's perspective	American Journal of Preventive Medicine	2007
Gonzalez-Rodriguez, Diego; Kostakis, Vasilis	Information literacy and peer-to-peer infrastructures: An autopoietic perspective	Telematics and Informatics	2015
Phoha, V. V.	An interactive dynamic model for integrating knowledge management methods and knowledge sharing technology in a traditional classroom	Book Chapter	2001
Schulz, Peter J.; Nakamoto, Kent	Patient behavior and the benefits of artificial intelligence: The perils of "dangerous" literacy and illusory patient empowerment	Patient Education and Counseling	2013
Durko, M.; Stoffova, V.	Perception: determinants and nature of direct and indirect experience	Book Chapter	2016
Robinson, L.; Bawden, D.	The story of data" A socio-technical approach to education for the data librarian role in the CityLIS library school at City, University of London	Library Management	2017
Troseth, Georgene L.; Strouse, Gabrielle A.; Russo Johnson, Colleen E.	Early Digital Literacy: Learning to Watch, Watching to Learn	Book Chapter	2017
Fagherazzi, G.; Ravaud, P.	Digital diabetes: Perspectives for diabetes prevention, management and research	Diabetes & Metabolism	2018
Gelles, Abby	Robotics and artificial intelligence as educational tools for developing self-sufficiency. Part I — Descriptive overview	Microprocessing and Microprogramming	1984
Dent, Valeda F.	Intelligent agent concepts in the modern library	Library Hi Tech	2007
Viana, Cassandra Lúcia de Maya	O impacto das inteligências artificiais na formação dos bibliotecários e cientistas da informação: revisão de literatura	Ciência da Informação	1990

Source: (The authors)

*Table 4. Theoretical groups of papers*

Research Group	Authors	Topic
Group I	Phoha (2013); Durko & Stoffova (2017); Troseth (2017); Strouse & Russo Johnson (2017), Dent (2007); Mathews et. al. (2003); Terracina & Mecella (2015).	Papers that propose computer models to learning mediation. They focus on digital literacy.
Group II	Raycheva (2018); Hesse & Shneiderman (2007); Schulz & Nakamoto (2013); Fagherazzi & Ravaud (2018)	Papers that study social, political and economical changes as a result of information technology. They analyze ethical and legal issues related to artificial intelligence.
Group III	Gonzalez-Rodriguez & Kostakis (2015)	Papers that focus on information literacy as a relevant factor to human development. The aspects of artificial intelligence help to develop flexible processes such as information retrieval and decision-making.

Source: (The authors)

interdisciplinary which results from postmodern science. The papers show three groups that differ theoretically, which is illustrated in Table 4 below.

Mathews *et. al.* (2003) and Terracina & Mecella (2015) suggest to adopt digital tutors and mentors to simulate discussions and tutoring strategies in the field of long-distance education. They propose that such mentors would act as teaching agents with the aim of motivating the students with respect to 21<sup>st</sup> century aspects of learning, such as autonomy, creativity, communication, critical thinking, and information literacy. These systems could guide the student's research by directing him or her to reliable and relevant sources of information.

This triggers that ethical issue that was mentioned above, the fear that technology will gradually replace human labor. On the other hand, it creates a new problem for the advancement of the learning that contemporary society requires. Morin (2003) states that this education demands people who can learn how to learn. This is based on the idea of two entities whose functions are clearly defined: (i) a mediator which stimulates learning by means of the effective and ethical use of existing informational resources and (ii) the learner who develops the necessary mental skills to build meaning.

This conclusion suggests that although algorithms and other elements of artificial intelligence are present in various modern activities, they can't replace the humanizing aspect of communication, which implies the formation of symbols by means of interlocution between thinking social beings. With that said, we can't ignore the fact that AI is very lucrative for long-distance education providers since they can hire fewer online tutors.

This research is concerned with individuals' lifelong learning since its objective is to contribute to "learning to learn" using artificial intelligence. Regarding the inference of categories, the articles point out the social consequences of technology and express a concern with the ethical evaluation of information and its sources. Hence, they correspond to the content analysis categories proposed by Bardin (2010).

The Group II research listed in Table 4 focuses on the social, political, and economic implications of the use of AI. They highlight the ethico-legal aspects linked to the use of AI, particularly for monitoring patient care. In those researches (Hesse; Shneiderman, 2007; Fagherazzi & Ravaud, 2018) the term "information literacy" only appears in a marginal way, even though those aspects are an integral part of the learning that is stimulated by various literacy programs.

Other articles criticized the use of AI in the medical field. There is an increasing production and sale of those advanced technologies, but physicians are not yet ready to deal with them. In the USA, for example, there have been cases of surgical complications due to the incorrect use of such tools.

The articles draw attention to the need to develop health literacy so that patients become aware of which practices and technologies are associated with this intense interest by capitalists of the medical industry (Schulz & Nakamoto, 2013). From a more optimistic point of view, AI has been a major asset in interpreting exam results and the large volume of data involved in diagnosing a disease.

Ethics must be applied by regulating the smart equipment which is used indiscriminately. According to Pellegrini and Vitorino (2018, p.128): "[...] the ethical aspect of information literacy is directly linked to the ethical and efficient use of information, and the information professional is responsible for its ethical use."

These articles fall into categories 1, 2, and 3 because they discuss the question of ethics directly. They also point out the negative aspects and the capitalist interests of large companies that want to sell this equipment at all costs. The articles highlight the relevance of ethical information access and the importance of evaluating the ethical quality of information. Quality information is collected in a critical way in the health field and for the creation of equipment and smart systems based on ethical principles.

The research that makes up Group III deal with information literacy as a fundamental skill to learn in the 21<sup>st</sup> century. These studies do not focus on the technologies themselves but rather the use of information. Gonzalez-Rodriguez and Kostakis (2015) found that the inherent complexity of the enormous amount of data and information requires heterogenous networks made up of both human and artificial agents to improve the recovery of information, the inference of knowledge, and decision-making.

According to the Italian philosopher Berardi (2015) society is a prisoner of social networks, and many people believe the simple-minded information shared there.

“[...] the legal and responsible use of information, based in the laws and norms that govern information use in each country, and the ethical principles of respect, justice, solidarity, and compromise, which lead to citizenship and to collective well-being” (Pellegrini; Vitorino, 2018, p.130).

The articles mainly focus on the third analytical category since they address the question of sharing results in an ethical way: considering the consequences of the product, respecting intellectual property, and developing strategies to avoid unethical use of the product. These inferences fall into the third category, which is related to the use of information in information literacy (Lau, 2008). Information literacy also encompasses the other analytical categories. The use of information depends on one's effective access and evaluation.

In general, the articles concern the ethical development of technology. The aim of the research is to improve people's quality of life and to build meaningful knowledge. Benasayag (UNESCO, 2018, p.15) concurs: “meaning is created by humans, not by machines”.

The ethical aspects referred to above must be described within the context of philosophy of information, which according to Floridi (2004) includes the need for new codes of conduct with respect to technology, as well as new laws and regulations that incorporate this new technological dynamic. An information literate person criticizes the context of AI in an ethical way and proposes new solutions to confront the challenges of that context.

## **Information Literacy Ethical Standards for Artificial Intelligence Use**

Table 5 illustrates the categories created from the IFLA standards of information literacy (Lau, 2007) and the basic principles of artificial intelligence created by The United Nations Educational, Scientific and Cultural Organization (UNESCO, 2018) and The Asimov Laws of Robotics (1942) in his science fiction book known as *Robot*. The basic principles and the laws can guide the ethical use of artificial intelligence.

IFLA (Lau, 2007) standards and indicators have been used by some researchers and organizations. They are flexible enough to be applied in practical contexts, so IFLA standards are a reference for information literacy in Table 5. The UNESCO (2018) document is based on specialists' interviews and international papers in the field of artificial intelligence. This is a relevant document to explain ethical aspects of artificial intelligence. Asimov's laws of robotics are also used as guidance for the categories, and some scientific research in the field of robotics has used them as ethical guidelines (Takahashi & Shimizu, 2014; Vadymovych, 2017; Butazzo, 2008).

Category 1 of ethical information access is connected to Standard 1 of information literacy (Lau, 2007). Furthermore, category 1 is linked to the law that artificial

*Table 5. Categories and indicators to analyze the results*

IFLA (Lau, 2007) Information Literacy Standards	Basic Principles of Artificial Intelligence (UNESCO, 2018)	Category and Indicators to Analyze the Results
<b>Access</b> → The user accesses information efficiently and effectively	A robot cannot hurt a human being or allow that someone suffers any harm (Asimov, 1983). The same thing occurs with artificial intelligence, because it cannot threaten human life (UNESCO, 2018).	<b>Category 1</b> Individuals access information in an ethical way <b>Indicators</b> <ul style="list-style-type: none"> <li>• Define information needs and its consequences to society</li> <li>• Seek information in the right places and respect copyrights</li> <li>• Develop seeking strategies ethically</li> </ul>
<b>Evaluation</b> → The user evaluates information in a critically	A robot must obey human orders except in cases where they hurt other people like Asimov (1983) First Principle. The UNESCO (2018, p. 39) proposes some questions: algorithms trained in human language acquire prejudices because of stereotypes of data which are part of daily life". The UNESCO is worried about the emergence of discriminatory, racist and hostile machines.	<b>Category 2</b> Individuals evaluate the quality and ethics of information critically. <b>Indicators</b> <ul style="list-style-type: none"> <li>• Interpret and evaluate information based on ethical principles</li> <li>• Consider the impacts of information to society</li> <li>• Filter information to decide its quality and relevance</li> </ul>
<b>Use</b> → The user applies information critically and precisely	A robot should protect its existence as long as it does not conflict to the First or Second Principle (Asimov, 1983). The systems and equipment of artificial intelligence need to be safe to avoid losing data and information. Furthermore, according to UNESCO (2018) our mission is to guide an international ethical debate about theses changes.	<b>Category 3</b> Individuals develop equipment and smart systems based on ethical principles <b>Indicators</b> <ul style="list-style-type: none"> <li>• Share the results ethically</li> <li>• Consider the consequences of a product and respect intellectual property</li> <li>• Develop strategies to avoid the use of products in non-ethical activities.</li> </ul>

Source: (The Authors)

intelligence cannot hurt anyone (UNESCO, 2018). This category helps to analyze if individuals define information needs and its consequences to society. In addition, it allows someone to identify if people seek information properly and if they respect intellectual property. Category 2, which evaluates the ethical quality of information, is connected to Standard 2 of information literacy (Lau, 2007). This category helps to fight against humans' prejudice in machines (UNESCO, 2018). Furthermore, category 2 identifies if individuals understand the consequences of information analysis and if they filter information based on relevance. The third category verifies

if individuals develop smart equipment and systems ethically. This category is connected to the third standard of information literacy (Lau, 2007) and the security of machines. The category helps to identify whether AI theory considers ethical principles. Furthermore, the third category contributes to an analysis of whether researchers share those technologies ethically and consider the consequences of AI to society. Researchers and product developers need to develop strategies to avoid the use of AI in non-ethical activities.

## **Some International Ethics Frameworks for Artificial Intelligence**

Europe was the first region to share ethical guidelines for artificial intelligence. The name of the document is Ethics Guidelines for Trustworthy AI which was published in April 2019 by the High-Level Expert Group on Artificial Intelligence (AI HLEG). This guideline inspired the Australian Ethical Principles for Artificial Intelligence as well.

Trustworthy artificial intelligence must be developed based on three components: it should be lawful, ethical and robust. These three dimensions helped to construct a framework for AI which is based on fundamental human rights (AI HLEG, 2019). This document considers privacy, transparency, safety and non-discrimination to be important aspects of the framework. Figure 1 below illustrates the framework.

The framework for trustworthy AI is useful to guide courses and organizations. The framework is based on three main aspects: Lawful AI, Ethical AI and Robust AI. Lawful AI depends on the country and region, so the European Commission decided not to discuss it in the document. Ethical AI is the ethical principles that can guide AI research, courses and development. Robust AI is the result of Lawful AI plus Ethical AI. Robust AI is focused on fairness, applicability, prevention of harm and the respect for human autonomy (AI HLEG, 2018).

The Australian government consulted some specialists in the artificial intelligence field to create a guide of ethical principles. These principles are illustrated by Table 6.

The Australian Ethical Principles for Artificial Intelligence was based on the Framework for Trustworthy AI (Figure 1). These principles have eight dimensions that guide professionals and organizations. The AI should be human-centered to value individuals, fair to avoid discriminating against certain people, private to protect data, transparent, contestable to receive feedback from the public and accountable.

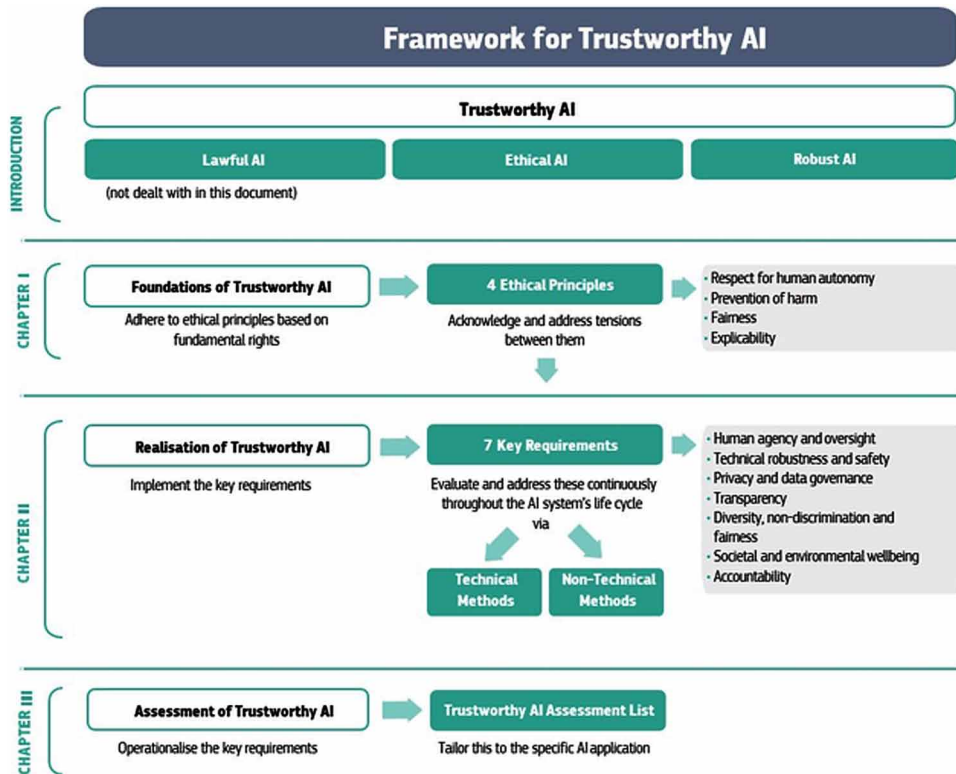
## **SOLUTIONS AND RECOMMENDATIONS**

Artificial intelligence is useful to solve some problems, and it can encourage the development of technology. Furthermore, AI makes individuals' work easier than in



Figure 1. Framework for trustworthy AI

Source: (AI HLEG, 2019, p. 8)



the last century. However, ethical and moral issues need to be evaluated by scientists and society. Some AI technology has been created to kill or to control people, and it can also be used to reduce freedom of speech. Individuals need to debate the limits of technology.

Information literacy can be applied in Computer Science, Analysis and Development of Systems, Artificial Intelligence and Data Science. This literacy can guide the ethical debate about the negative consequences to humans and the environment. Therefore, information literacy must be applied in the syllabus of courses and it should also be taught as a course. The aim is to graduate critical thinkers, so that they access information ethically.

The first step to fight against unethical AI is to understand the difference between the concepts of human and machine intelligence. The second step is to study artificial intelligence to help people. Scientists and research organizations can ask public opinion about the limits of technology, so that decisions can be made democratically and no one social group prevails.

*Table 6. Australian ethical principles for artificial intelligence*

Principles	Explanation
Human, social and environmental wellbeing	Throughout their lifecycle, AI systems should benefit individuals, society and the environment.
Human-centred values	Throughout their lifecycle, AI systems should respect human rights, diversity, and the autonomy of individuals.
Fairness	Throughout their lifecycle, AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
Privacy protection and security	Throughout their lifecycle, AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.
Reliability and safety	Throughout their lifecycle, AI systems should reliably operate in accordance with their intended purpose.
Transparency and explainability	There should be transparency and responsible disclosure to ensure people know when they are being significantly impacted by an AI system, and can find out when an AI system is engaging with them.
Contestability	When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or output of the AI system.
Accountability	Those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

Source (The Australian Government - Department of Industry, Innovation and Science)

## FUTURE RESEARCH DIRECTIONS

Researchers from several fields have a common mission: they need to raise awareness of the consequences and the problems of lack of ethics in innovation and artificial intelligence. According to the authors in the discussion section, machines and software can replace an ethical professional at work. However, employers need to consider that machines do not yet have empathy towards others.

The code and functions of technology must be analyzed ethically before being released to the public. Future research can apply information literacy in course curricula in Computer Science. Technology courses must consider the ethical implications of machines and systems. These technologies are helpful to people and must respect human rights instead of threatening them.

International organizations have created documents to share ethical code. The documents must be used to emphasize the relevance of information literacy and critical thinking to artificial intelligence development. An international movement and public policy are necessary to value this literacy and lifelong learning.

## CONCLUSION

These days, technology is present in all aspects of people's lives. On the one hand, there is intense optimism about the importance of Information Communication Technology, but on the other hand, several problems arise, including ethical ones related to their use. Artificial intelligence has transformed traditional ways of learning, so it demands that people become information literate. This literacy allows effective decision-making through critical thinking.

The ethical dimension of information literacy is fundamental to evaluate the consequences of artificial intelligence. Technological research has advanced quickly. However, the debate about the results of these technologies is still slow.

The systematic literature review demonstrates that there is not a lot of papers about information literacy and artificial intelligence. However, a few papers about both topics are available internationally. These papers are focused on digital tutors to help students learn online.

## ACKNOWLEDGMENT

This research was supported by the *Coordenação de Aperfeiçoamento de Nível Superior* (CAPES) and Conselho Nacional Científico e Tecnológico (CNPq) and *Fonds de Recherche du Québec – Nature et Technologie* (FRQNT).

## REFERENCES

- Adolphs, P., & Epple, U. (2015). *Status Report: Reference Architecture Model Industrie 4.0 (RAMI4.0)*. Retrieved from: [https://www.zvei.org/fileadmin/user\\_upload/Themen/Industrie\\_4.0/Das\\_Referenzarchitekturmodell\\_RAMI\\_4.0\\_und\\_die\\_Industrie\\_4.0-Komponente/pdf/5305\\_Publikation\\_GMA\\_Status\\_Report\\_ZVEI\\_Reference\\_Architecture\\_Model.pdf](https://www.zvei.org/fileadmin/user_upload/Themen/Industrie_4.0/Das_Referenzarchitekturmodell_RAMI_4.0_und_die_Industrie_4.0-Komponente/pdf/5305_Publikation_GMA_Status_Report_ZVEI_Reference_Architecture_Model.pdf)
- Almada-Lobo, F. (2015). The Industry 4.0 revolution and the future of Manufacturing Execution Systems (MES). *Journal of Innovation Management*, 3(4), 16–21. doi:10.24840/2183-0606\_003.004\_0003
- Association of College & Research Libraries. (2000). *Information Literacy Competency Standards for Higher Education*. Retrieved from: <https://alair.ala.org/handle/11213/7668>

Australian Ethical Principles for Artificial Intelligence. (2019). *The Australian Government - Department of Industry, Innovation and Science*. Retrieved from: <https://consult.industry.gov.au/strategic-policy/artificial-intelligence-ethics-framework/>

Belluzzo, R. C. B. (2014). O conhecimento, as redes e a competência em informação (CoInfo) na sociedade contemporânea: Uma proposta de articulação conceitual [Knowledge, network and information literacy (IL) in current society: a conceptual discussion]. *Perspectivas em Gestão & Conhecimento, João Pessoa, 4*, 48–63.

Bordeleau, A. F., Mosconi, E., & Santa-Eulalia, L. A. (2018). Business Intelligence in Industry 4.0: State of the art and research opportunities. *Proceedings of the 51st Hawaii International Conference on System Sciences*. Retrieved from: <http://hdl.handle.net/10125/50383>

Bostron, N. (2016). *Artificial intelligence: 'we're like children playing with a bomb'*. Retrieved from: <https://www.theguardian.com/technology/2016/jun/12/nick-bostron-artificial-intelligence-machine>

Demasson, A., Partridge, H., & Bruce, C. (2016). Information literacy and the serious leisure participant: Variation in the experience of using information to learn. *Information Research, 21*(2).

Ethical Guideliness for Trustworthy Artificial Intelligence. (2019). *European Comission. High-Level Expert Group on Artificial Intelligence (AIHLEG)*. Retrieved from: <https://ec.europa.eu/futurium/en/ai-alliance-consultation>

Floridi, L. (2004). Open problems in the philosophy of information. *Metaphilosophy, 35*(4), 2004. doi:10.1111/j.1467-9973.2004.00336.x

Grafstein, A. (2017). Information Literacy and Critical Thinking: Context and Practice. Pathways into Information Literacy and Communities of Practice: Teaching Approaches and Case Studies, 3-28.

Lau, J. (2007). *Guidelines on Information Literacy for Lifelong Learning*. The Hague: IFLA. Retrieved from: <https://www.ifla.org/files/assets/information-literacy/publications/ifla-guidelines-en.pdf>

Lloyd, A. (2007). Recasting information literacy as sociocultural practice: Implications for library and information science researchers. *Information Research, 12*(4).

Mccarthy, J. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence, 1955*. Retrieved from: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

Obama, B. (2009). *National information literacy awareness month*. Retrieved from: <https://www.govinfo.gov/app/content/pkg/STATUTE-123/pdf/STATUTE-123-Pg3711.pdf>

Otonicar, S. L. C., Valentim, M. L. P., & Feres, G. G. (2015). Competência em informação e os contextos educacional, tecnológico, político e organizacional [Information literacy and the educational, technological, political and organizational contexts]. *Revista Ibero-americana de Ciência da Informação, Brasília*, 9(1), 24–142.

Schwab, K. (2016). *The fourth industrial revolution*. Crown Business.

Slayton, R. (2018). *Policy Series: Beyond Cyber-Threats: The Technopolitics of Vulnerability*. Retrieved from: <https://issforum.org/roundtables/policy/1-5bc-technopolitics>

The Australian Government - Department of Industry. (2019). *Innovation and Science. AI Ethics Principles*. Retrieved from: <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>

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

United Nations Educational, Scientific and Cultural Organization (UNESCO). (n.d.). Towards a global code of ethics for artificial intelligence research. In *Artificial Intelligence: the promises and the threats*. Retrieved from: <https://en.unesco.org/courier/2018-3/towards-global-code-ethics-artificial-intelligence-research>

Vitorino, E. V., & Piantola, D. (2011). Dimensões da competência informacional (2) [Dimensions of information literacy]. *Ciência da Informação, Brasília*, 40(1), 99–110.

Yafushi, C. A. P. (2015). *A Competência em informação para a construção de conhecimento no processo decisório: estudo de caso na Duratex de Agudos (SP)* [Information literacy to construct knowledge in the decision-making process: a case study at Duratex Agudos (SP)] (Master's Dissertation). Retrieved from: <https://repositorio.unesp.br/handle/11449/126599>

## ADDITIONAL READING

- Capurro, R. (2006). Towards an ontological foundation of information ethics. *Ethics and Information Technology*, 8(4), 175–186. doi:10.1007/10676-006-9108-0
- Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, 20(1), 1–3. doi:10.1007/10676-018-9450-z
- Gomes de Andrade, N., Pawson, D., Muriello, D., Donahue, L., & Guadagno, J. (2018). Ethics and Artificial Intelligence: Suicide Prevention on Facebook. *Philosophy & Technology*, 31(4), 669–684. doi:10.1007/13347-018-0336-0
- Lloyd, A. (2017). Information literacy and literacies of information: A mid-range theory and model. *Journal of Information Literacy*, 11(1), 91–105. doi:10.11645/11.1.2185
- Pellegrini, E., & Vitorino, E. (2018). A Dimensão Ética da Competência em Informação sob a Perspectiva da Filosofia. *Perspectivas em Ciência da Informação*, 23(2), 117–133. <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2953>. doi:10.1590/1981-5344/2953
- Russel, S. (2015). Robotics: Ethics of artificial intelligence. *Nature*, 521(7553), 415–418. doi:10.1038/521415a PMID:26017428
- Silva, H., Jambeiro, O., Lima, J., & Brandão, M. A. (2005). Inclusão digital e educação para a competência informacional: Uma questão de ética e cidadania. *Ciência da Informação*, 34(1), 28–36. doi:10.1590/S0100-19652005000100004
- Steinerová, J., & Ondříšová, M. (2019). Research Data Literacy Perception and Practices in the Information Environment. In *Information Literacy in Everyday Life. ECIL 2018. Communications in Computer and Information Science* (Vol. 989). Springer. doi:10.1007/978-3-030-13472-3\_51


## KEY TERMS AND DEFINITIONS

- Artificial Intelligence:** The intelligence of technology and machines.
- Critical Thinking:** The ability to criticize a text or information.
- Ethics:** Part of philosophy that studies human values.
- Information Evaluation:** It is the analysis of information quality.
- Information Literacy:** Ability to construct knowledge and interpret information.
- Information Science:** Scientific field that studies information in different contexts.
- Interdisciplinarity:** The knowledge produced by two or more scientific field.

# Chapter 11

## Artificial Intelligence and Ethical Marketing

**Bistra Konstantinova Vassileva**

 <https://orcid.org/0000-0002-5976-6807>  
*University of Economics, Varna, Bulgaria*

**Plamena Palamarova**

*University of Economics, Varna, Bulgaria*

### ABSTRACT

*In this chapter, the author argues that technologies will transform the marketing organization and reshape the marketing activities of companies. The aim of the chapter is to summarize the main challenges of digital disruption as well as to identify their implications to the legal and ethical aspects of digital and interactive marketing activities. The research aims driving this chapter are related to the identification of the main challenges regarding the legal protection of social media customers. Survey results about social media behaviour and cybersecurity issues are presented and discussed.*

### INTRODUCTION

The dynamics of information and communication flows over the last decade have not only changed radically, but unlike previous decades, a significant amount of information is ‘produced’ and received outside organizations. In terms of the cost of information processing, for organizations and end users it is more efficient to use ICT in an increasing number of situations. What a decade ago was a very limited vertical flow of information and communication is now a vast sea of vertical and

DOI: 10.4018/978-1-7998-4285-9.ch011

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

horizontal flows ‘pushed’ and ‘pulled’ by an incredible variety of offline and online sources. Market interactions and relationships are so intense and transformative that it is difficult to distinguish between market players and participants in communication space due to the overlap of the economic domains of consumers, competitors and collaborators. The convergence between the clearly defined economic sectors of audiovisual media, telecommunications and information and computer technology years ago poses many challenges to marketers, including the legal and ethical aspects of the use of new media.

## BACKGROUND

According to the results of the global digital business survey conducted by MIT Sloan Management Review and Deloitte’s, the implementation of digital technologies by organizations in their activities depends on the level of maturity of the digital business. Companies in transition to digital maturity focus on the changes in the process of integration of digital technologies (social networks, mobile technologies, analytics and cloud technologies), and the ultimate goal of integration is to transform the way their business is organized and operated. These companies are characterized by transformed processes, commitment of human resources and flexible business models. As the level of digital maturity increases, companies are beginning to develop and implement the above four digital technologies to almost the same extent. The main goals of companies with a lower level of digital maturity are related to solving specific business problems using individual digital technologies.

*Table 1. Classification of companies at the stage of Web 4.0*

Criterion	Class 1	Class 2	Class 3	Class 4
Digital maturity (DQ)	Established leaders	Emerging leaders	Followers	
Digital strategy	Smaller-scale disrupters	Fast-followers	Digital relocators	Business models reshapers
Level of digital transformation	Pure-play global industry disrupters	Ecosystems shapers	Incumbents	

Source: Adapted by: van Bommel, E., Edelman, D. and Ungerman, K. Digitizing the consumer decision journey. McKinsey Company, 2014; Visser, J., Field, D., and Sheerin, A. The Agile Marketing Organization, The Boston Consulting Group, October 2015.



Irrespective of their reactions to digital disruption all organizations are forced to change their business routine. The complex changes caused by turbulent markets, aggressive global competition, the rapid emergence of new technologies and destructive innovation are the reason for moving to the next phase in the evolution of marketing - Marketing 4.0, which is defined as an extremely fast cybernetic system of incentives, feedback and proactive reactions with a focus on flexible processes and detailed knowledge of the business (Table 2).

*Table 2. Key characteristics of Marketing 4.0*

Characteristic	Description
Timing	<ul style="list-style-type: none"> <li>§ Speeding up marketing activities</li> <li>§ Continuous adaptation</li> <li>§ Scrum approach to plan and implement marketing activities</li> </ul>
Talant management	<ul style="list-style-type: none"> <li>§ Qualified analysts</li> <li>§ Data experts</li> <li>§ Customer experience officer</li> <li>§ Content officer</li> <li>§ Data storyteller</li> <li>§ Data scientists</li> <li>§ Multi-channel campaign manager</li> </ul>
Data and analytics	<ul style="list-style-type: none"> <li>§ Data management</li> <li>§ Advanced analytics</li> <li>§ Consumer insights</li> </ul>
Marketing organization models	<ul style="list-style-type: none"> <li>§ Degree of centralization</li> <li>§ Focus: products, segments, channels, geography, function</li> </ul>

Source: Adapted by Vassileva (2017)

The establishment of the marketing organization as a cybernetic system allows real-time monitoring of global transactions and consumer activities in the market space, which, in turn, requires the use of new approaches to the marketing processes organization and activities and their legal regulation. Customers are placed at the center of this new digitally-based marketing system. The system elements and their relationships should be precisely planned to stimulate customers' interactions with the products, to offer customers emotional personal experience (through the so called 'touch points') and to add value. Networks could be used as illustrative example for such a system where nodes (vertices) are the elements of the digital marketing mix and 'Organisation2Customer' interactions are the edge (links) (Vassileva, 2017).

## **MAIN FOCUS OF THE CHAPTER**

### **Social Media and the Legal Challenges Facing Digital Marketers**

Social media is most often defined as a group of Internet-based applications that are built on the theoretical and technological foundations of Web 2.0, allowing the creation and exchange of user-generated content (Kaplan and Haenlein, 2010). At the heart of this definition is the ability of social media to facilitate interaction between users, and in this sense they include blogs, content communities, discussion boards and chat rooms, sites for products and/or services reviewing, virtual worlds and social networks. Social networks include applications such as Facebook and Twitter, which allow users to construct personal profiles and compile a list of people with whom to share and interact (Boyd and Ellison, 2007). Internet-based communication, especially through social media, is becoming a key factor influencing various aspects of consumer behavior and purchasing decisions. Digitalization transforms the purchasing decision-making process, including the way customers search for information, view and evaluate products and services, interact with the organization, and make purchases. Traditional purchasing decisions have been transformed as a process into ‘digital consumer journey’ (van Bommel et al., 2014). The phenomenon of social media leads to the formation of a unique social media ecosystem, consisting of a diverse combination of social media sites that vary in terms of their scope and functionality (Kietzmann et al., 2011:241). Social networks form an essential part of the modern social media ecosystem.

Social media has become an essential part of consumers’ daily lives, attracting the attention of marketers who have included them as a key tool in today’s hybrid marketing mix. According to Kaplan and Haenlein (2010), the current ‘social media revolution’ is transforming the Internet back to its roots as a platform aimed at facilitating the exchange of information between users. Social media offers an unprecedented opportunity for companies to search for and analyze behavioral patterns of consumer communication and interaction on social networks so that they can develop appropriate marketing activities, especially brand marketing communication campaigns, CRM, e-mail marketing, etc. Such increasing dependence on social media raises various challenges. On the one hand, marketers should apply appropriate personal data protection and should take into account various ethical issues related to the idea of individuality, i.e. that each of us has a unique core identity, and that we’re not interchangeable. This situation reveals the importance of protection of customers’ freedom not to be predictable and protection of customers’ privacy (Bughin et al., 2018). The topic of ethical issues related to the use of social media and other means of digital marketing comes under the focus of normative theories,

which aim to develop a set of good practices to guide people's behavior. In the field of digital marketing, this is related to the theoretical statements about the information transmitted through social media, according to which people can control the end result of the application of the technology and on this basis to build control systems, incl. for preventive activities against unethical activities and unethical behavior. On the other hand, there is no legal protection and no remedy against the adverse effects of any changes in the social network's policies or loss of investment due to such changes (Härtig Attorneys at Law, 2013). Companies should always have a plan B for different scenarios and be prepared with for the worst case scenario in case of any failures of the network, especially those companies that rely solely on social platforms for their corporate and marketing communications.

The main area of research in the field of legal regulation of digital and interactive marketing is the impact of technology and intervention policies on social media. Some basic areas of intervention are shown in Table 3.

*Table 3. Basic areas of social media interventions*

Areas of Intervention	Legal Regulation
Personal data protection: § online (incl. identity theft) § corporate data surveillance § contextual integrity <sup>1</sup>	§ Partial § Significant differences between countries
Interruption or suspension of support for social networking services	Lack of effective legally regulated activities
Changes in the rules of use of social networks	Lack of effective legally regulated activities
Use of the trademark in social media and in domain registration	§ Significant differences between countries § The online space is out of the scope of intellectual property protection
Identification of companies in company websites	Significant differences between countries
Responsibility for content and content management, incl. posts	Lack of effective copyright protection
'Hidden' advertising	Significant differences between countries

Source: Author's work

Dataveillance (a combination of 'data' and 'surveillance') is a new word for the misuse of personal data in the online space, first introduced by Clarke (1991). The main problems of users regarding protection of their personal data are related to the illegal use of these data for commercial purposes as well as their illegal transfer to third parties (contextual integrity). Unfortunately, users do not know who collects their personal information, what kind of information is collected, why it is collected

*Table 4. Selected privacy and data protection cases*

Company	Activity	Date	Information Source
Yahoo	Secretly scanned customer emails for U.S. intelligence	October, 2016	<a href="https://www.reuters.com/article/us-yahoo-nsa-exclusive-idUSKCN1241YT">https://www.reuters.com/article/us-yahoo-nsa-exclusive-idUSKCN1241YT</a>
Facebook	Germany bans Facebook on WhatsApp users' data gathering	September, 2016	<a href="https://bit.ly/37oANcY">https://bit.ly/37oANcY</a>
Google LLC	The DPA breaches for the Safari Workaround	October, 2018	<a href="https://bit.ly/3odz0gI">https://bit.ly/3odz0gI</a> <a href="https://www.judiciary.uk/wp-content/uploads/2018/10/lloyd-v-google-judgment.pdf">https://www.judiciary.uk/wp-content/uploads/2018/10/lloyd-v-google-judgment.pdf</a>
Google Inc	NT1 and NT2 data protection and privacy claims brought in respect of Google's refusal to delist/deindex search results in two spent conviction cases	February, 2018	<a href="https://bit.ly/3odIrgc">https://bit.ly/3odIrgc</a>

Source: Author's work

Note: DPA – Data Protection Act

and how it is collected. Much of this process is automated, which further reduces the power of users to influence and/or control their own personal information.

The problem of online security is extremely important for marketers. Having too many security options and settings confuses users. Many of these alternatives are quite broadly defined, e.g. friends. In this regard, fewer but easier-to-understand options would significantly improve consumers' protection options. The problem of providing the choice to determine the area of security (especially with regard to personal data) is left in the hands of users, ie. companies outsource this activity to their customers. The implementation of legal regulation in this field is still widely discussed. Some the 'hottest' topics include the following: use of facial recognition software and the legal framework to determine its lawfulness, the application of the territorial scope of the right to be forgotten to worldwide domain names, the application of reasonable expectation of privacy in the context of the "voyeurism" offence, misuse of private information and copyright infringement in social media and any digital environment, etc.

An inherent threat from social media is that information users are 'lulled to sleep' by the false sense of security from websites that use personal information for profit. Many social networks are still not profitable and this increases the pressure on them to provide access to personal information to their users, especially by advertisers who want to target potential users personally according to their interests, incl. Development of their psychographic profiles.

## Social Media and Cybersecurity

The results of conducted online survey among Bulgarian students show that cyber security is associated predominantly with the security of personal data, information security and safe use of Internet (Table 5). There are skeptic respondents who think that cyber security is not possible to be ensured and that any cyber protection could be breached.

*Table 5. Attitudes toward cyber security, %*

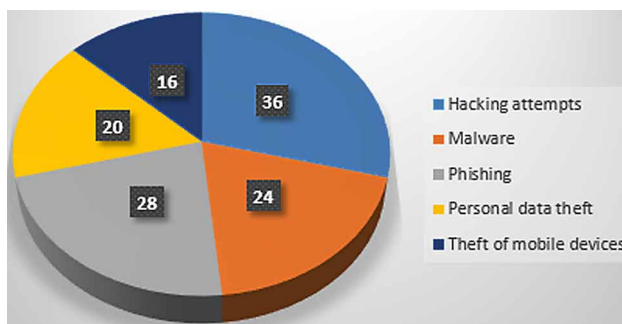
Statements	Agree	Neither Agree, nor Disagree	Disagree
1. Information security for citizens and companies should be an extraordinary priority of the state.	88	4	8
2. Bulgarian companies are provided with enough legal tools to feel safe about the security of the information they work with.	44	20	36
3. The misuse of internal information should be subject of criminal law.	72	20	8
4. Employees who work with sensitive internal information are vulnerable for accusation of its misuse.	64	32	4
5. Signing confidentiality statements cannot limit the misuse of internal firm information.	84	12	4

Note: Standard Likert scale is used. Data are presented in percentages.

Approximately 80% of the respondents cannot mention any technology or technical tool which can be used for cyber security purposes. Among the indicated tools are cryptography, antivirus software, digital signature.

*Figure 1. Attack types that were faced by respondents last year, %*

Source: Author's work

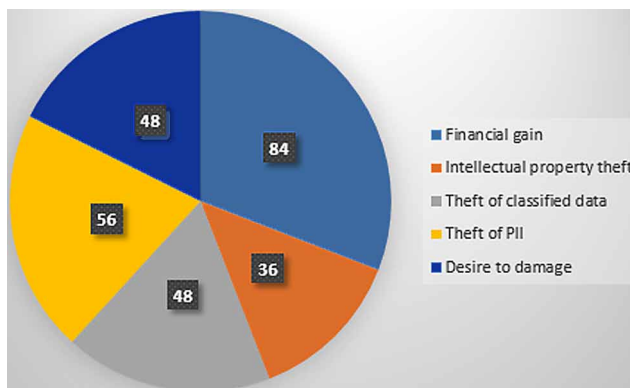


More than 80% of respondents believe that cyber criminals are driven by financial motivation, i.e. they were paid for their cyber attacks (Figure 2).

*Figure 2. Driving forces of cyber threats*

*Source: Author's work*

*Note: Respondents were instructed to indicate the three most important motives; PII – Personally Identifiable Information*



The desire to acquire personal and sensitive information is indicated as a second by importance driving motive. On the third place are emotional motives, e.g. desire to harm and boredom. At the moment of the survey, the misuse of intellectual property rights is considered the least important motive for cyber attacks.

## **SOLUTIONS AND RECOMMENDATIONS**

The dramatic paradigm shift provoked by the shift in control of communications from companies to customers raises serious challenges to marketers. The most consistent characteristic of social media is the participatory, two-way, decentralised accommodation of communication (Rader et al., 2015:194-5). Companies should reconsider their traditional marketing strategies and tactics because social media are not merely an additional available channel of communication. Given the raising interest in social media, companies have realized that their presence and performance in these environments requires rules, regulations and instructions of use, in order to effectively and consistently communicate the company's policies regarding social media participation (Fuduric and Mandelli, 2014:8). Companies should focus on the following key legal and ethical issues regarding their social media strategies. First, personal and individual-related issues which are related with

eventual disclosure of confidential information, defamation issues, privacy. Second, intellectual property rights and correlated issues of unauthorized use of trademarks, unauthorized use of copyright-protected works. Third, business-related issues such as human resources issues (hiring and retention practices based on social media reviews), employees' misuse or improper use of social media in terms of removing unfavorable information or adding untrue unfavorable information, use of fake identities to sabotage competitors, etc.

## **FUTURE RESEARCH DIRECTIONS**

Future research has to take a critical look at the differences on the micro-level of everyday consumer practices between various consumers and consumer groups, in order to assess the "privacy divide" in an everyday surveillance environment (Pierson and Heyman, 2011:39). This includes investigation and analysis not only of the knowledge of customers about privacy issues (awareness), but also their capabilities (what customers are able to do), their attitudes (incl. their preferences), and their practices (what customers actually do). Moreover, research should focus not only on demand side of the market (customers) but on supply side as well (companies) and the stakeholders (incl. legislation, institutions, governments, etc.). Companies must be aware of the potential pitfalls when using social media and performing their digital marketing activities which calls for 'clear' legal requirements across countries and markets.

## **CONCLUSION**

Today, modern communications are inextricably linked to the development of digital devices and wireless transmission of information. The development of digital devices and communications in general has led to an increase in the speed of information processing, an increase in the speed of communications and access to them, as well as to a reduction in the prices of digital communications. The digitalization of communications leads to improved efficiency, mass and wide availability of easy and fast ways of communication between people.

The speed of development of digital communications and technologies poses many challenges in terms of legal regulation of online activities, both for businesses and for institutions and all stakeholders. Some of the opportunities to overcome the problems with social media can be sought in the following directions. Social media developers can use new forms of media protocols, which are defined as smart or adaptive. The choice of the most appropriate social network design can be

done on the basis of the ‘infosphere’. An unstructured centralized architecture is considered to be more efficient than the standard one in terms of control. Last but not least, there is a need of target investment of resources by the government and higher education institutions to raise young people’s awareness of the consequences of inappropriate and irresponsible use of information provided by social media, as well as cyber security.

## REFERENCES

- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x
- Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes From the AI Frontier: Modeling the Impact of AI on the World Economy*. McKinsey Global Institute. Available at: <https://mck.co/3aIkDZT>
- Clarke, R. (1991). Information technology and dataveillance. In C. Dunlop & R. Kling (Eds.), *Controversies in Computing*. Academic Press.
- Dholakia, N., Zwick, D., & Denegri-Knott, J. (2010). Technology, Consumers, and Marketing Theory. In *The SAGE Handbook of Marketing Theory* (pp. 494–511). SAGE.
- Flatow, I. (2008). *Web privacy concerns prompt Facebook changes*. In *On Science Friday*. NPR ScienceFriday Inc.
- Fuduric, M., & Mandelli, A. (2014). Communicating social media policies: Evaluation of current practices. *Journal of Communication Management (London)*, 18(2), 158–175. doi:10.1108/JCOM-06-2012-0045
- Härting Attorneys at Law. (2013). *Legal aspects of social media*. Available at: [https://www.haerting.de/sites/default/files/downloads/handout\\_legal\\_aspects\\_of\\_social\\_media\\_2013.pdf](https://www.haerting.de/sites/default/files/downloads/handout_legal_aspects_of_social_media_2013.pdf)
- Introna, L. D. (2007). Maintaining the reversibility of folding: Making the ethics (politics) of information technology visible. *Ethics and Information Technology*, 9(1), 11–25. doi:10.1007/10676-006-9133-z
- Jara, A. J., Parra, M. C., & Skarmeta, A. F. (2012). Marketing 4.0: A New Value Added to the Marketing through the Internet of Things. *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, 2012 Sixth International Conference on, 852-857.



Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review*. Available at: <<https://sloanreview.mit.edu/digital2015>>

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59–68. doi:10.1016/j.bushor.2009.09.003

Kietzmann, J., Hermkens, K., McCarthy, I., & Silvestre, B. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), 241–251. doi:10.1016/j.bushor.2011.01.005

Mangold, W., & Faulds, D. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, 52(4), 357–365. doi:10.1016/j.bushor.2009.03.002

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review (Seattle, Wash.)*, 79, 101–139.

Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the ‘IT’ in IT research – a call to theorizing the IT artifact. *Information Systems Research*, 12(2), 121–134. doi:10.1287/isre.12.2.121.9700

Pierson, J., & Heyman, R. (2011). Social media and cookies: Challenges for online privacy. *Info*, 13(6), 30–42. doi:10.1108/14636691111174243

Rader, C., Subhan, Z., Lanier, C., Brooksbank, R., Yankah, S., & Spears, K. (2014). CyberRx. *International Journal of Pharmaceutical and Healthcare Marketing*, 8(2), 193–225. doi:10.1108/IJPHM-05-2013-0027

Solove, D. J. (2004). *The Digital Person: Technology and Privacy in the Information Age*. University Press.

van Bommel, E., Edelman, D., & Ungerman, K. (2014). *Digitizing the consumer decision journey*. McKinsey Company. Available at: <http://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/digitizing-the-consumer-decision-journey>

Vassileva, B. (2017). Marketing 4.0: how technologies transform marketing organisation. *Obuda University e-Bulletin*, 7(1), 47–56.

Visser, J., Field, D., & Sheerin, A. (2015). *The Agile Marketing Organization*. The Boston Consulting Group. Available at: <https://www.bcg.com/publications/2015/marketing-brand-strategy-agile-marketing-organization>

## **ENDNOTE**

- <sup>1</sup> Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79, 101-39.

## Chapter 12

# Liability in Labor Legislation: New Challenges Related to the Use of Artificial Intelligence

**Andriyana Andreeva**

*University of Economics, Varna, Bulgaria*

**Galina Yolova**

*University of Economics, Varna, Bulgaria*

### ABSTRACT

*The study analyzes the influence of artificial intelligence on labor relations and the related need to adapt to the legal institute of liability in labor law with the new social realities. The sources at European level are studied and the current aspects of liability in the labor law at a national level are analyzed. Based on the analysis, the challenges are outlined and the trends for the doctrine, the European community, and the legislation for the introduction of a regulatory framework are identified.*

### INTRODUCTION

Responsibility is a social phenomenon and a category that accompanies the development of human society. The types of responsibility: moral, ethical, political, legal, etc. have arisen and followed the development of the state and society. In accordance with the historical and social stages, the various means used by the states to regulate the ongoing processes have also changed. Legal liability is one of the types of responsibility, and it is a guarantee for the implementation of legal orders. Each branch of law is related to a specific type of legal liability, which corresponds to its peculiarities and is adequate to the degree of public danger of

DOI: 10.4018/978-1-7998-4285-9.ch012

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

offenses (Dimitrova,D,Mateeva,Zv,Dimitrova,D, 2020, pp. 108-119). At the same time, legal liability is associated with the development of public relations and, respectively, the overall evolution of the legal theory and the legal system. On the one hand, each stage presupposes that the legal institutes correspond to the needs of the specific time and actually cover and regulate both the positive and negative processes with the necessary means. However, the introduction of new concepts, necessarily following the social development, presupposes the timely adaptation of the existing institutions. In its original form, legal liability has been adapted to influence individuals who have committed offenses in the different spheres of law. Subsequently, it has expanded its perimeter, and currently all known legal entities can be its addressees, as per the relevant legal branch, and according to the specific varieties of legal liability.

The importance of the topic of liability as a legal concept is indisputable, as it is embedded in and forms an integral part of the law, and along with the specifics of the legal departments it is associated with basic concepts such as offense, sanction, subjective rights, etc.

Modern society has entered a new dimension of its development, which is associated with the fourth industrial revolution. It is reflected in all spheres of society, but also provokes a need for reaction on the part of the states, institutions at the international and national levels, as well as on the part of the civil society. In the hands of socially irresponsible companies, digitalisation can lead to huge profits at the expense of circumventing the rules. (Blagoycheva, 2020, p. 55). Law as a whole is facing a time of cardinal changes, requiring rethinking and upgrading of legal concepts (Andreeva,A,Yolova,G., 2020, pp. 11-18) (Andreeva,A,Yolova,G., 2020, pp. 22-35), (Andreeva, A., Yolova, G., Dimitrova, D., 2020, pp. 43-48) (Andreeva, A., Yolova, G., 2019, стр. 178-188) (Andreeva, A., Yolova, G., Dimitrova, D., 2019, стр. 44-47) (Andreeva,A,Yolova,G., 2018, стр. 320-328) (Andreeva,A,Yolova,G., 2018, стр. 293-307) supplementing legal institutes, as well as introduction of new principles of public relations regulation. Among these issues is also the topic of liability when using artificial intelligence.

As already mentioned, legal liability is a complex concept that comprises different types, some of which are typical for a certain branch, for example, civil liability in the civil and commercial law, criminal liability in criminal law, administrative criminal liability in administrative law (Tsankov, P., Andreeva, A., Yolova, G., Dimitrova, D, 2006, стр. 124-127), etc.

Labor law does not stipulate a liability that covers the offenses in their totality, ie there is no complex and unified type of “labor liability” applicable to the different types of violations of labor law. (Mrachkov, 2013, стр. 11).

On the other hand, three main types of liability are applicable in labor law - disciplinary, property and administrative penal liability. These three types of legal

liability are designed to create the necessary guarantee for compliance with labor law, and according to the specifics of the offenses, they impact in a specific way and with different intensity (Andreeva,A,Yolova,G., 2011). The first type (disciplinary liability) (Andreeva,A,Yolova,G., 2020)) is the most common for this legal branch in which liability actually arose as a type. For the other two types models are used which are provided in the civil and administrative law respectively, but have been adapted to the needs of labor law.

These three types of liability are applicable with respect to subjects in the categories of worker, employee or employer, respectively. In the hypotheses of labor law violation they can be applied independently or cumulatively.

The introduction of artificial intelligence in the work process requires the transformation of liability and its rethinking in a number of aspects. On the one hand is the need to update the traditional types of legal liability applicable in this legal branch, both with regard to the offences hypotheses and in terms of the sanction applied to the responsible subjects.

On the other hand, if we regard this concept in a broader sense, a special type of liability should be stipulated for using artificial intelligence, respectively, for damage caused by robots replacing the traditional workforce and endangering jobs, which is also evident in the intensified efforts for urgent creation of a unified regulatory framework on European and international level. In these particular instances, however, we are not talking about the classical concept of ‘liability’ and it would therefore not be included in the research scope of the authors of this study.

In terms of subject matter, the authors have set a number of restrictions, examining as a matter of priority the substantive side of liability and not going into procedural issues. At the same time, the study aims to identify trends and raise such topical issues on the doctrine field that currently do not have a specific normative expression or are still in the process of debate in the institutions at European level.

## **BACKGROUND**

**The topicality of the researched topic, concerning the transformation of liability in labor law in accordance with the new tendencies related to the introduction of artificial intelligence** is indisputable and provokes debate, both at the scientific level and in international and EU institutions. Aspects of AI and their impact on public relations are yet to be explored and analyzed. In the present study, the issue is raised in terms of the impact of AI on the classical legal institute of liability, and the authors have tried to go beyond traditionalism and give a new perspective based on a philosophical idea, through the prism of the new realities.

**The aim of this study** is to research the need to adapt the legal institute of liability in law to the new social realities and the introduction of AI in the labor process, both by studying the trends of common European documents and by arguing hypotheses for its adequate and timely reflection in the national legislation.

To achieve this aim, the authors have performed the **following tasks**:

1. Studying the specifics and trends for the regulation of the liability in the use of AI in the common European framework;
2. Studying the relevant aspects of liability in the current condition and the variety of types in Bulgaria's labor law;
3. Outlining trends following the introduction of AI in public relations and guidelines for creating a regulatory liability mechanism.

**The methodological basis** of this study is related to the cumulative use of traditional methods for legal research such as: comparative-legal, formal logical-legal and general scientific methods of knowledge: induction, deduction, analysis and synthesis. The study is in accordance with national legislation and leading European sources as at July 30, 2020.

## **LIABILITY IN THE USE OF ARTIFICIAL INTELLIGENCE – COMMON EUROPEAN FRAMEWORK AND PRINCIPLES**

The increasing processes of digitalization and the intensity of their influence on public relations are happening at a pace that the legal framework can hardly react to or anticipate. This is a trend with a disturbing influence on the concepts of legal doctrine, ethical and legal philosophy, and, of course, the normative mechanism.

Although there exist a stable framework for ensuring product safety, some European documents, such as the Product Liability Directive and the Machinery Directive, are only the basis that needs to be upgraded, however with qualitatively different regulatory tools. For example, the Product Liability Directive provides that if a defective product causes any damage to consumers or their property, the manufacturer should provide compensation, notwithstanding whether due to negligence or fault on his part. However, this provision should undoubtedly undergo transformations, and in some cases it cannot even be applied, given the qualitatively different nature of digital technologies, already featuring neuromorphic chips, world-class high-performance computers, as well as leading quantum technology projects for human brain mapping.

The processes of activating the legislative mechanism are a leading European policy of priority, appealing for a sustainable and urgent, but also comprehensive

and unified regulatory framework. Intensified processes for debating the issue of AI use are taking place at various expert and institutional levels, but the fact is that there is still no comprehensive Community framework for defining liability for AI use, which is increasingly perceived as a trend with a worrying impact on social and ethical aspects of public relations.

We can summarize that the core of the debate focuses on **three main aspects, namely:**

- the extent to which decision-making autonomy extends in the boundary between human and artificial intelligence,
- who should be responsible for decision-making by autonomous cyber-physical systems in traditional labor relations, until recently dominated by human capacity only,
- and last but not least, what should be the nature and type of liability of artificial intelligence users.

These three fundamental issues are entirely within the moral and social case philosophy and the lasting influence of the ethical norm on the specific legally regulated relations.

At the same time, the issue should not lag behind the fourth basic aspect of liability, **namely the social responsibility upon displacing the traditional workforce**. In this sense, the desire to create a regulatory framework should not only outline the specifics of liability for the use of AI and the ensuing damages, but to anticipate or in parallel outline the hypotheses concerning the replacement of human labor with all subsequent social, moral and institutional problems.

The new vision of a digital Europe, based on the Strategy for the Digital Future of the European Union and the Strategy for a Single Data Market, undoubtedly links the success of the digital transition to the requirement that new technologies be used in the spirit of traditions in order to strengthen democratic civil and social values. In this sense, all three key objectives in creating a common European legislation in the digital field are repeatedly outlined in the following directions, namely: technologies that work for people, a fair and competitive economy, an open, democratic and sustainable society.

The limited nature of the study does not ensure a detailed system in the analysis of the various institutional acts, and therefore we should limit ourselves to the aspects of only the most important and crucial AI-regulating acts.

The decisive development of the wide-ranging debate stems from the **European Parliament's Motion for a Resolution containing recommendations to the Commission on civil law for robotics (2015/2103 (INL))**, based on a report dated 27.01.2017. The resolution concerns the use of autonomous machines, with

reflexive potential consequences in contractual and non-contractual liability, and in this sense it is considered necessary to clarify the liability for the actions of robots, the legal capacity and / or legal status of robots and artificial intelligence in order to ensure transparency and legal certainty for producers and consumers across the European Union. In particular, this concerns autonomous vehicles, care robots, medical robots, human “repair” and upgrading, and unmanned aerial vehicles.

The Liability section of the document highlights the urgent Community issues that directly affect the new paradigm of responsibility in the different light addressed by the traditional legal doctrine and the philosophy of law. Based on the understanding that robot autonomy can be defined as the ability to make decisions and execute them in the outside world, regardless of external control or influence, they generally set out the unconditional framework of the principles from which the subsequent regulatory mechanisms are to unfold. We can summarize these frameworks and issues in the following more important directions:

- whether the currently developed and enforced liability rules are sufficient, respectively whether (and this is indisputable) new principles and rules for the legal liability of the various participants for the actions and omissions of the robots are required;
- the extent to which the autonomy of robots presupposes a permanent revision of the existing legal categories, respectively, whether a new category should be stipulated, with its own specific characteristics and consequences;
- the applicable and effective legal framework does not imply liability for damages caused by robots’ actions or omissions vis-à-vis third parties;
- in a situation where a robot can make autonomous decisions, the traditional rules cannot be applied, as they would not allow the person responsible for providing benefits to be specified,
- there is an urgent need for up-to-date rules, in line with technological developments, not only to overcome the shortcomings of the current legal framework in the field of contractual liability, but also to introduce qualitatively different aspects, insofar as they concern machines designed to choose their contractors, to enter into contracts and make decisions;
- the current legal framework is not sufficient to cover damage caused by the new generation of robots, as they can be equipped with adaptive and learning capabilities leading to a certain degree of unpredictability of their actions, as these robots would learn on their own from their own variable experience and would interact with their environment in a unique and unpredictable way;
- non-contractual liability applied under Council Directive 85/374 / EEC can only cover damage caused by manufacturing defects of a robot, hence the argument that the framework of objective, guiltless liability may be



insufficient. Based primarily on the need for a normatively established permanent definition of the concept of AI in a single regulatory framework, the motions for resolutions focus on **three main axes**: principles for the use of autonomous cyber physical systems, including the introduction of a system for registration of advanced robots based on the criteria for classifying robots, a regulatory framework based on ethical aspects, in particular the robotics charter and ethical rules for AI users, and last but not least, of course, general liability frameworks.

The proposals for the development of the framework liability in the subsequent institutional acts are at the levels of general civil liability based on objective liability, or risk management approach, compulsory insurance schemes based on the manufacturer's obligation to take out insurance for the autonomous robots it produces, adequate insurance system through a fund to guarantee compensation in cases of missing insurance coverage, compensation fund. In any event, the emphasis is on the rule that "the legal solution applicable to the liability of robots and artificial intelligence in cases other than property damage should in no way limit the type or extent of the damage that can be recovered, nor to limit the forms of compensation that may be offered to the party concerned solely on the grounds that the damage was caused by an agent other than a human being."

At the same time, the main goal of the proposal for a single charter of robotics is the need for compliance with ethical standards on the part of researchers, specialists, users and designers, as well as for the introduction of a procedure to enable the resolution of relevant ethical dilemmas and to enable these systems to operate in an ethically responsible manner. This is further developed by the proposals for specific ethical principles of developers and users of AI, in particular inclusion, accountability, security, reversibility, confidentiality of personal data and respect for fundamental rights, maximizing benefits while minimizing possible harm.

The latter is especially relevant in the field of labor relations. In the latter case, there is a reasonable emphasis on the understanding that "usually the risk of harm should not be greater than in normal life, ie people should not be exposed to risks that are greater than or in addition to those to which they are exposed in their normal way of life. The operation of the robotics system should always be based on an in-depth risk assessment process, which should be based on the principles of precautionary measures and proportionality."

As a follow-up to the Proposal, in its Communication on **Artificial Intelligence for Europe to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, the Commission presented a first definition of AI - COM (2018) 237 final** the Commission assumes that the term 'artificial intelligence' (AI) is used for systems that exhibit intelligent behavior

by analyzing their environment and, with some degree of autonomy, taking action to achieve specific goals. AI-based systems can be fully software - operating in the virtual world (eg voice assistants, image analysis software, search engines, voice and face recognition systems), or can be implemented in hardware devices (eg advanced robots, autonomous cars, drones or Internet applications of objects).

This definition has been refined by the High Level Expert Group, clarifying that “artificial intelligence (AI) systems are software (and possibly hardware) systems created by people who, in view of a complex purpose, operate within the physical or the digital dimension, perceiving their environment by collecting data, interpreting the collected structured or unstructured data, substantiating on the basis of knowledge or processing the information obtained from this data, and deciding on the best action (actions) to achieve the given goal. AI systems can either use symbolic rules or learn a digital model, and can adapt their behavior by analyzing the way the environment has been affected by their previous actions.”

The overall essence of the Community philosophy is summarized by the objective set out in the Communication, namely to ensure an appropriate ethical and legal framework, which is based on the values of the Union and is in line with the EU Charter of Fundamental Rights. This is deemed necessary, linked to the establishment of regulatory frameworks on existing product liability rules, detailed analysis of emerging challenges, cooperation with stakeholders through the European AI Alliance, with a view to developing guidelines on AI-related ethical issues.

At the same time, it clearly outlines the two main key challenges facing the Union, namely the need to increase the digital skills capacity of workers and, at the same time, a high level of involvement of social systems in job transformation by ensuring access to social protection for all citizens, including workers and the self - employed, in accordance with the European Pillar of Social Rights.

A whole new debate on a strong legislative response was launched with the **White Paper on AI - Europe in search of excellence and a climate of trust (Brussels, 19.2.2020 COM (2020) 65 final) and the accompanying report on the safety and accountability framework**. The envisaged follow-up framework at Community and national level focuses on a future regulatory framework for AI in Europe, which will create a unique “trust ecosystem” based on an anthropocentric, trust-based approach to human-oriented AI. Thus, the principles formulated in the document are aimed at different options for the use of autonomous robots, but with a common goal - fairness, non-discrimination, reasonable measures to ensure that the use of AI systems will not lead to results that cause prohibited discrimination.

Closely focused on formulating the specifics of the property levels: contractual or non-contractual liability for damage caused by autonomous systems, and primarily addressing the defined risk areas and areas of use, the White Paper establishes principles of reflection also in the employment relations, which should urgently be

adapted both in the further development of the Community framework and upon introduction into the national legislation. At the same time, an important emphasis of the document is the understanding that “in view of its (the AI’s – authors’ note) importance for individuals and for the *acquis communautaire* in the area of employment equality, the use of AI applications in the recruitment process, as well as in situations affecting workers’ rights, it will always be considered “high risk” and therefore the following requirements will apply by default”.

When formulating the principles of liability, the emphasis is also on the types of mandatory legal requirements for participants in the processes of using AI at the following levels: training data, data and documentation, information on provision, reliability and accuracy, human supervision, and specific requirements for some specific AI applications, such as those used for remote biometric identification.

Obviously, this is an argument for the need to formulate a legal framework in which increased human intervention at a certain stage should ensure as much as possible the absence of full autonomy in the decision-making algorithm of machines with the goal of ensuring minimum social, but also purely ethical human rights.

It is clear that at this stage the debate still raises more questions than answers, and these questions are a serious challenge not only to the regulatory mechanism, which is expected to adequately respond to unfamiliar instruments of intervention in social relations. They are a serious problem for the legal philosophy and doctrine, requiring urgent reform and rethinking of classical legal institutions, related to both legal liability and the subject of law, and in particular - to the holder of the right to work.

## **THE NEED TO ADAPT LIABILITY IN LABOR LAW WITH THE INTRODUCTION OF AI IN EMPLOYMENT RELATIONS**

It is more than indisputable that labor law is an important area requiring a revision of the rules, in which the introduction of AI is particularly manifested. This branch of law in the national legal system of each country is a compilation combining both current international norms and national norms - a specific reflection of the ongoing processes in the country. At the current stage of development of the industry, specific decisions related to guarantees of compliance with labor law are the responsibility of national legislation on the one hand, as individual countries are at different levels of introduction of technological processes and in particular of robotics and automation, which duplicate and / or replace the traditional forms of labor. At the same time, and in view of the globalization process, national institutions are part of the international community, which is called upon to study the ongoing processes

regardless of national borders and to develop mechanisms for the protection of the labor of all workers and employees.

AI and robots have long entered different countries at different levels. Technologically developed countries such as Japan, China, USA, etc. have stipulated in their legal systems norms regulating labor relations and in general the use of AI for various purposes. The process of application of these new technologies is developing at a pace that cannot be decided only by individual countries, but requires the unification of the international community in the development of standards and principles applicable in international relations.

The challenge is to create a set of legal guarantees that will preserve man as the highest value of public relations and put human beings in a role guaranteeing them an equal start in the labor market in competition with the new intelligent systems.

The problem of liability in labor law is complex and requires a combination not only of traditional models of legal liability and their updating with new sanctions, expansion of the subjects, but also the provision of new principles of liability.

This fully applies to European legislation, which currently remains fully applicable, despite the fact that AI is included in the processes, and needs to be supplemented and updated. The issue, which is also studied at European level and is the subject of analysis in national institutions and doctrine, is the level of adequacy of the legal framework and the existing legal institutions to the risks posed by AI systems.

In order to refine the regulation concerning liability in its diversity of types within the labor process, it is necessary to **identify the main risks**. In this way, the composition of offenses of a certain type could be formed and adequate corrections could be sought in the legislative framework at European and national level.

The first place can be occupied by the risk concerning the compliance with the legal framework guaranteeing the fundamental rights of the subjects, respectively subjective labor rights of the workers and employees. This risk is linked to the lack of comprehensive AI legislation covering both its characteristics and the legal principles in its use. In this sense, there is a “gap” in the legislation, which makes it difficult for law enforcement institutions to develop hypotheses of offenses involving AI systems and respectively seeking accountability. **The main difficulties are at several levels: related to the detection and proof of offenses, as well as to the identification of the subject of liability.**

In this aspect it can be said that there is a need for urgent correction in the institute of liability, both in its general theoretical aspect and in view of the types of legal liability.

Next in the list of risks should be the safety of the products including AI. European legislation contains certain rules concerning the safety of products, but priority is given to the launch of such products. There is no regulation for the liability of the subjects in the implementation and inclusion of products / systems with AI in the

labor process. In this case, in addition to the traditional types of legal liability of labor law, the application of a complementary type should be provided, regulating the liability between the employer who introduced AI in a certain labor process and the manufacturer / supplier / developer of the product with AI.

Equality in the field of employment and maintaining the central role of individuals in the labor process should be considered as a leading priority in the regulation of liability. In this regard, when introducing AI systems in the labor process situations in which there is a danger to the rights of workers, respectively for their privacy, should be perceived as risky.

As a prevention measure in the labor process, it is advisable to consider the introduction of requirements for the parties to the employment relationship in the future legislative framework. In the first place, they should focus on the employer, since in his capacity as an entity with an employer's legal capacity, he is responsible for the overall organization of the process. The employer is also responsible for the effective management and integration of the organization of safe and healthy work in a given organization (Blagoycheva, Andreeva & Yolova, 2019). Given the specifics of the individual types of work, detailing could be done in the special regulations governing the relevant field, such as health care, transport, etc.

**The following components** could be included in the requirements for employers when implementing AI systems:

- Information about the manufacturer and supplier;
  - Reliability and accuracy;
  - Mechanisms for internal control over the labor process;
  - Complementing the control with human supervision over the work of the systems with AI;
  - Obligations of employers ensuring that privacy and personal data are adequately protected when using products and services with AI. **Respectively, requirements to the workers and employees** participating in work processes involving AI;
  - Mandatory initial and ongoing training to work with AI systems;
  - Compliance with safety rules and instructions for working with AI systems.
- The current Bulgarian legislation lacks regulations concerning the inclusion of AI in the work process, respectively the commitments of the parties to the employment relationship and regulation of specifics in liability for violations, although in the field of information and communication technologies (ICT) a number of regulations are established and operative in Bulgaria. The most significant of these are the Personal Data Protection Act (Promulgated SG No. 1 of January 4, 2002, amended and supplemented SG No. 17 of February 26, 2019), the Electronic Communications Act (Promulgated SG No. 41

of 22 May 2007, amended and supplemented, SG No. 17 of 26 February 2019), Law on Electronic Government (Promulgated, SG No. 46 of 12 June 2007, amended and supplemented SG No. 94 of 13 November 2018), Law on Electronic Document and Electronic Certification Services (Promulgated SG No. 34 of 6 April 2001, amended SG No. 1 of 3 January 2019), Law on the Commercial Register and the Register of Non-Profit Legal Entities (Promulgated SG No. 34 of April 25, 2006, amended and supplemented, SG No. 38 of May 10, 2019), Law on Provision of Distance Financial Services (Promulgated SG No. 105 of December 22, 2006, amended SG No. 20 of March 6, 2018), Cyber Security Act (Promulgated SG No. 94 of November 13, 2018). At the same time, the development and application of information technologies is subject to many strategic documents: National Program “Digital Bulgaria 2025” and the Roadmap for the period 2018 - 2025 of the Ministry of Transport, Information Technology and Communications, Concept for digital transformation of Bulgarian industry (Industry 4.0), National Strategy for Cyber Security, “Cyber Sustainable Bulgaria” 2020, Innovation Strategy for Smart Specialization of the Republic of Bulgaria 2014 - 2020, etc.

We believe that the steps that should be taken are **on a number of levels:**

- to introduce basic requirements in the general source - the Labor Code, which should supplement the content of the employment relationship with rights and obligations when using AI in the labor process;
- to regulate in the Labor Code the offenses committed in a work process with implemented AI systems;
- to expand the hypotheses of legal liability. This applies to all three types - disciplinary, property and administrative liability;
- to regulate the liability involving the persons in the chain developer-supplier-implementer

user employer. The subject of labor law are the labor relations and the related other public relations (art. 1, para 1 of the Labor Code). Within the subject of this legal branch is also the need to ensure the observance of the norms through the type diversity of three legal responsibilities which are different in their nature. After the occurrence of the employment relationship and within its existence, the worker and employee provide their labor force for the benefit of the employer, however the latter undertakes a commitment to organize the labor process in a way that preserves the life and health of workers and employees and to minimize the risks related to the specifics of the respective activity. Each work process includes risks that should

be properly identified and the necessary and adequate measures for ensuring safe and healthy working conditions should be taken. The use of AI systems is a new hypothesis for Bulgarian employers, since the technological processes have not yet been widely introduced in all areas.

Material liability which is regulated in Chapter Ten of the Labor Code occupies an important place among the legal liabilities applicable in labor legislation. It contains two variants from the entity's point of view, namely 'Employer's Liability' in Section I and 'Employee's Liability' in Section II. This symmetry in material liability is not accidental, it is stipulated by the legislator in order to correspond to the obligations of the parties to the employment relationship and in this sense it represents liability for compensation for damages caused by failure to fulfill obligations on the part of the respective party (Articles 126, 127, 275 Labor Code). Alternatively, we can share the views expressed in the legal doctrine that the material liability of one party to the other is their secondary legal obligation arising from their main obligations. (Кожухаров, 1958, стр. 314) (Варкалю, 1978, стр. 26-29). Therefore we believe that in view of the material liability of the parties to the employment relationship in cases of implementation of AI systems in the labor process, first of all it is necessary to include in the sources norms regulating the rights and obligations of the subjects with regard to AI. This should be regulated both by state sources at the level of the general act – the Labor Code, and by special laws in order to reflect the peculiarities of the specific activity and the level of participation of intelligent systems. Next, the employer should be obliged to supplement its internal acts with obligations which should reflect the specifics of the use of AI in the respective unit: enterprise, institution, etc. On this basis material liability, which is also symmetric, could subsequently occur for the parties, in view of the synalagmatic and bilateral nature of the employment relationship.

There is no explicit text in the Bulgarian legislation to commit the parties to any liability in the presence of implemented AI systems, which does not mean that no liability is borne; however, its definition as a type would make it difficult for the enforcers because of the lack of practice. We believe that in the future, when updating the legislation in order to guarantee the rights of the parties to the employment relationship, the material liability of the employer should be expanded. Hypotheses for liability of the employer must be explicitly provided in both cases: firstly for damages caused to an employee due to an accident at work or occupational disease, when this is due to the use of AI, and secondly, material liability for violation of employees' rights, supplementing the exhaustive list in Chapter Ten with one or more specific rights linked to and arising from the use of AI. AI has repeatedly been the subject of research interest in the legal literature, with the authors focusing on various aspects thereof (Angusheva, 1987, стр. 53-68) (Vasilev, 1997, стр. 215) (Vasilev, Obezsheteniya po trudovoto pravootnoshenie, 2012, стр. 526-542) (Sredkova,

2011, стр. 387-396) (Balabanov, 1988, стр. 29-35) (Balabanov B., 1990, стр. 16-19) (Vasilev, *Za pravната sashtnost na imushtestvenata otgovornost mezhdu stranite*, 1993,2, стр. 31-41), as well as a comprehensive study (Mrachkov, 2013, стр. 17) .

In its essence, material liability is civil and in this sense we join the statements in legal theory advocating this position. (Goleva, 2011, стр. 15-16). The functions of the employer's material liability, namely compensatory, stimulating, restorative and protective (Mrachkov, 2013, стр. 30-33), correspond to the need to address the issues of inclusion of AI in the labor process, however with maximum safeguards for employees. Therefore, we believe that this liability is the first of the three types of liabilities of labor law, which should be transformed to cover the hypotheses of damage caused by the use of AI systems.

Respectively, consideration should be given to supplementing the liability under labor law with the general civil liability, and after the realization of the material liability of the employer the latter should have the opportunity to file a claim against the supplier or the manufacturer of the AI systems.

The present study cannot cover all possible nuances of the transformation of liability in labor law with the inclusion of AI in the labor process and the occurrence of legal offenses. It aims to raise the main issues that arise at this stage and to seek logic in reasoning at the doctrinal level, in order to support the future legal framework in this direction.

## **FUTURE RESEARCH DIRECTIONS**

The advent of AI into the field of labour relations is still unpredictable as a process in terms of its benefits and harms. On the one hand, it is indisputable that AI affects technological processes at a hitherto unseen pace, but on the other hand it entails the permanent displacement of human workforce, requires constant adaptability on the part of the workers through lifelong learning and enhancement of their digital skills and competencies, and poses serious ethical problems, especially in the use of autonomous cyber-physical systems capable of making independent decisions without human intervention. It is precisely in this regard that there is an instant need for urgent measures to adapt the regime of liability in labour law.

It is evident that there have been intensified actions at the European institutions level calling for a timely adoption of a regulatory framework defining the type, principles and limits of liability. The qualitative solution of these problems lays the basis for subsequent adaptation of the national legislations. In this sense, without denying the need for adjustments in labour legislation, it is clear that basic and stable legal concepts should undoubtedly be revised, concerning traditional hypotheses



of liability in the labour process, whose adjustment could prove to be a serious and socially significant problem.

## **CONCLUSION**

The process of adapting the liability in labour law at national level should not lag behind the trends emerging from the debates at Community level. At the same time it is clear that this is a process that requires a rethinking of a number of traditional and enduring legal concepts. In this respect, the authors believe that adjustments should be made in several different aspects in order to cover individual legal concepts and constructs and bring these up to date comprehensively so as to achieve maximum applicability. The proposals of the authors in this regard concern:

- introduction of new rights and obligations of the parties, given their role as the basis of a guaranteed labour process involving AI systems;
- introduction of provisions for new offenses in the different types of liability in labour law, including typical hypotheses of acts related to AI;
- further development of the characteristics of the different types of liability in labour law and their expansion to cover cases of offenses related to the use of AI.

Traditional labour law is not yet at the stage of rethinking the hypotheses about the liability of the employer in using AI capable of autonomous decision-making. Although this is a serious debate within European Community institutions, it remains outside the overall national legal doctrine and its concepts of the person liable. In this sense, this is a problem that has yet to be addressed. At the same time, the problem should not be avoided and the emphasis should always be on the fact that labour legislation, based on the principles of equal and non-discriminatory rights, the most important of which is the right to work, should uphold the principle of shared social responsibility in the process of replacement of the traditional labour force with AI.

## **REFERENCES**

Andreeva, A., & Yolova, G. (2011). *Yuridicheska otgovornost i kontrol za spazvane na trudovoto i osiguritelno zakonodatelstvo*. Varna: Univ. izd. Nauka i iekonomika, Bibl. Prof. Tsani Kalyandzhiev.

Andreeva, A., & Yolova, G. (2018). Predizvikatelstva i tendentsii pred sotsialnata zashtita v usloviyata na digitalното obshtestvo. *Izvestia Sp. Ikonomicheski universitet - Varna, Varna: Nauka i ikonomika*, 62, 3.

Andreeva, A., & Yolova, G. (2018). *Za subekt na pravoto na trud i predizvikatelstvata na tehnologichното obshtestvo. Tsifrova ikonomika i blokcheyn tehnologii: Edinadeseta mezhdunarodna nauchno-prilozhna konferentsia, 29.06 - 01.07.2018g.: Sbornik nauchni trudove*. Varna: Largo siti.

Andreeva, A., Yolova, G. (2019). The Challenges of the Fourth Industrial Revolution Faced by the Labour Market: European and National Processes and Trends. In *Mezhdunarodni klasterni politiki: Balgaro-kitayski forum: Sbornik s dokladi ot mezhdunarodna konferentsia*. Varna: Nauka i ikonomika.

Andreeva, A., & Yolova, G. (2020a). Trudovopravnite printsipi - evolyutsia i transformatsia v erata na digitalizatsia i izpolzvaneto na izkustven intelekt. *Izvestia. Sp. Ikonomicheski universitet - Varna, Varna: Nauka i ikonomika*, 64, 22-35.

Andreeva, A., & Yolova, G. (2020b). Transformatsia na pravната vrazka rabotodatel – rabotnik v rezultat na vliyanieto na digitalizatsiyata. *De Jure, V. Tarnovo*, 11, 11-18.

Andreeva, A., & Yolova, G. (2020c). *Trudovo i osiguritelno pravo. Vtoro preraboteno i dopalнено izdanie*. Varna: Nauka i ikonomika.

Andreeva, A., Yolova, G., & Dimitrova, D. (2019). Artificial intellect: Regulatory Framework and Challenges Facing the Labour Market. In *ompSysTech '19: 20-th International Conference on Computer Systems and Technologies, 21 - 22 June 2019, University of Ruse: Proceeding*. New York: ACM Digital Library.

Andreeva, A., Yolova, G., & Dimitrova, D. (2020). Computer Technology and Ehealth. Trends and Regulatory Framework. *Economics and Law, Blagoevgrad: South-West Univ. Neofit Rilski Publ. House*, 2, 43–48.

Angusheva, V. (1987). Pravna uredba na otgovornostta na predpriyatieto za trudova zlopoluka i profesionalno zabolyavane v novia Kodeks na truda. *Pravna misal*, 33, 53-68.

Balabanov, B. (1988). *Novi momenti v imushtestvenata otgovornost mezhdu predpriyatieto i rabotnika po Kodeksa na truda*. DP, 6.

Balabanov, B. (1990). *Osnovaniето na iska pri smart ot trudova zlopoluka (chl.200 KT i chl.49 ZZD)*. DP,6.

Blagoycheva, H. (2020). The digitization as a stimulus for corporate social responsibility. In CSR and Socially Responsible Investing Strategies in Transitioning and Emerging Economies. IGI Global. doi:10.4018/978-1-7998-2193-9.ch003

Blagoycheva, H., Andreeva, A., & Yolova, G. (2019). Obligation and Responsibility of Employers to Provide Health and Safety at Work – Principles, Current Regulation and Prospects. *Economic Studies*, 28(2), 115–137.

Byalata kniga za II - Evropa v tarsene na visoki postizhenia i atmosfera na doverie (Bryuksel, 19.2.2020 COM (2020) 65 final)

Dimitrova, D., Mateeva, Z., & Dimitrova, D. (2020). *Administrativno pravo i protses*. Varna: Nauka i ikonomika.

Goleva, P. (2011). *Deliktno pravo*. Feneya.

Izkustven intelekt za Evropa COM/2018/237 final Saobshtenie na komisiyata do evropeyskiaparlament,saveta,ikonomiceskia i sotsialenkomitet i komitetanaregionite. <https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=CELEX%3A52018DC0237>

Kozhuharov, A. (1958). *Obligatsionno pravo. Obshto uchenie za obligatsionnoto otnoshenie*. Sofia: Nauka i izkustvo.

Mrachkov, V. (2013). *Imushtestvena otgovornost na rabotodatelya*. Sibi.

Predlozhenie za rezolyutsia na evropeyskia parlament sadarzhashto preporaki kam Komisiyata otnosno grazhdanskopravni normi za robotikata (2015/2103(INL)). [https://www.europarl.europa.eu/doceo/document/A-8-2017-0005\\_BG.html#title1](https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_BG.html#title1)

Sredkova, K. (2011). *Trudovo pravo. Spetsialna chast. Dyal I. Individualno trudovo pravo*. Sofia: UI Sv. Kl. Ohridski.

Tsankov, P., Andreeva, A., Yolova, G., Dimitrova, D. (2006). *Osnovi na publichnoto pravo*. Varna: Nauka i ikonomika, IU-Varna.

Varkallo, V. (1978). *Ob otvetstvenosti po grazhdanskomu pravu*. Progres.

Vasilev, A. (1993). *Za pravната sashtnost na imushtestvenata otgovornost mezhdu stranite*. Zakon.

Vasilev, A. (1997). Imushtestvvena otgovornost na rabotodatelya za vredi ot trudova zlopoluka i profesionalno zabolyavane. *Interyus*, 215.

Vasilev, A. (2012). *Obezhtetenia po trudovoto pravootnoshenie*. Sibi.

## ADDITIONAL READING

Mrachkov, V. (2010). *Trudovo pravo*. Sibi.

Mrachkov, V., Sredkova, K., & Vasilev, A. (2009). *Komentar na Kodeksa na truda*. Sibi.

Sredkova, K. (2011). *Trudovo pravo*. Sv. Kliment Ohridski.

## KEY TERMS AND DEFINITIONS

**AI:** A concept first introduced at the Dartmouth Summer Research Project on Artificial Intelligence (1956). It is generally accepted for this concept to be used as a designation of a branch of science and a research discipline. Currently, the definition relates to the functionality of software systems performing actions by analysing experience and data.

**Damage:** An adverse consequence suffered as a result of unlawful conduct – violation or crime committed by another person. The law prescribes that the damage contains the elements losses sustained and loss of profit. The main division of damages is into material and non-material (non-pecuniary) damages.

**Disciplinary Liability:** A separate type of liability, applicable in parallel with the other types of liability provided for in the criminal, administrative and civil law. Disciplinary liability is prescribed as a sanction for the culpable unlawful conduct of the employee in connection with violation of the labour discipline. It consists in the imposition of the disciplinary sanctions stipulated in the Labour Code and in the obligation of the employee to suffer certain adverse consequences.

**EU Legal Acts:** Sources of norms for the Community legal order. They are mainly divided into primary legislation, including the Treaty on European Union, the Treaty on the Functioning of the European Union and the protocols thereto, the Charter of Fundamental Rights of the European Union, the Treaty establishing the European Atomic Energy Community (Euratom), international agreements and general principles of EU law. The secondary legislation includes regulations, directives, decisions, recommendations, and opinions which the institutions of the Union may adopt only if express competence has been conferred on them by the provisions of the Treaties.

**Liability for Damages:** Liability arising as a consequence of an unlawful act or non-performance of a contractual obligation, hence its division into contractual and tortious liability. Tortious liability presupposes existence of the elements damage, act, unlawfulness of the act, guilt and causal link between the act and the damage.

**Liability in Labour Law:** A complex concept, which encompasses the types of legal liability applicable in labour law: disciplinary, material and administrative-penal liability. It is used for the needs of the legal doctrine, but there is no legally defined uniform term.

**Material Liability:** In the doctrine of labour law and for the purposes of the present study, this is considered as a means of compensating the damages that a party to the employment relationship suffers as a result of the culpable unlawful conduct<sup>1</sup> of the other party. An exception to this is the material liability of the employer for damages suffered by the employee as a result of an accident at work or an occupational disease (Article 220 of the Labour Code), which is not a result of the conduct of the employer.

## ENDNOTE

- <sup>1</sup> An exception to this is the material liability of the employer for damages suffered by the employee as a result of an accident at work or an occupational disease (Article 220 of the Labour Code), which is not a result of the employer's conduct.

## Compilation of References

Abawajy, J. (2014). User preferences of cyber security awareness delivery methods. *Behavior & Information Technology*, 33(3), 236-247.

Adolphs, P., & Epple, U. (2015). *Status Report: Reference Architecture Model Industrie 4.0 (RAMI4.0)*. Retrieved from: [https://www.zvei.org/fileadmin/user\\_upload/Themen/Industrie\\_4.0/Das\\_Referenzarchitekturmodell\\_RAMI\\_4.0\\_und\\_die\\_Industrie\\_4.0-Komponente/pdf/5305\\_Publikation\\_GMA\\_Status\\_Report\\_ZVEI\\_Reference\\_Architecture\\_Model.pdf](https://www.zvei.org/fileadmin/user_upload/Themen/Industrie_4.0/Das_Referenzarchitekturmodell_RAMI_4.0_und_die_Industrie_4.0-Komponente/pdf/5305_Publikation_GMA_Status_Report_ZVEI_Reference_Architecture_Model.pdf)

Agger, I., & Jensen, S. (1996). *Trauma and Healing Under State Terrorism*. ZEB Books.

Agievich, V. (2014). *Mathematical model and multi-criteria analysis of designing large-scale enterprise roadmap*. PhD thesis.

Al-Daeef, M. M., Basir, N., & Saudi, M. M. (2017, July). Security awareness training: A review. In *Proceedings of the World Congress on Engineering (Vol. 1, pp. 5-7)*. Academic Press.

Alderman, L. (2019). French Court Fines UBS \$4.2 Billion for Helping Clients Evade Taxes. *The New York Times*. Retrieved October 2019, from <https://www.nytimes.com/2019/02/20/business/ubs-france-tax-evasion.html>

Al-Janabi, S., & Al-Shourbaji, I. (2016). A study of cyber security awareness in educational environment in the middle east. *Journal of Information & Knowledge Management*, 15(1). doi:10.1142/S0219649216500076

Almada-Lobo, F. (2015). The Industry 4.0 revolution and the future of Manufacturing Execution Systems (MES). *Journal of Innovation Management*, 3(4), 16–21. doi:10.24840/2183-0606\_003.004\_0003

AMInfo. (2014). *Middle Eastern clients in the HSBC Switzerland leaks*. Swiss Leaks. Retrieved June 2017, from <http://ameinfo.com/luxury-lifestyle/list-middle-eastern-clients-in-the-hsbc-switzerland-leaks/>

Anderson, R. (2008). *Security Engineering: A Guide to Building Dependable Distributed Systems* (2nd ed.). Wiley Publishing, Inc.

Andreeva, A., & Yolova, G. (2011). *Yuridicheska otgovornost i kontrol za spazvane na trudovoto i osiguritelno zakonodatelstvo*. Varna: Univ. izd. Nauka i iкономика, Bibl. Prof. Tsani Kalyandzhiev.

- Andreeva, A., & Yolova, G. (2018). Predizvikatelstva i tendentsii pred sotsialnata zashtita v usloviyata na digitalното obshtestvo. *Izvestia Sp. Ikonomicheski universitet - Varna, Varna: Nauka i ikonomika*, 62, 3.
- Andreeva, A., & Yolova, G. (2018). *Za subekta na pravoto na trud i predizvikatelstvata na tehnologichното obshtestvo. Tsifrova ikonomika i blokcheyn tehnologii: Edinadeseta mezhdunarodna nauchno-prilozhna konferentsia, 29.06 - 01.07.2018g.: Sbornik nauchni trudove*. Varna: Largo siti.
- Andreeva, A., & Yolova, G. (2020a). Trudovopravnite printsipi - evolyutsia i transformatsia v erata na digitalizatsia i izpolzvaneto na izkustven intelekt. *Izvestia. Sp. Ikonomicheski universitet - Varna, Varna: Nauka i ikonomika*, 64, 22-35.
- Andreeva, A., & Yolova, G. (2020b). Transformatsia na pravната vrazka rabotodatel – rabotnik v rezultat na vliyaniето na digitalizatsiyata. *De Jure, V. Tarnovo*, 11, 11-18.
- Andreeva, A., & Yolova, G. (2020c). *Trudovo i osiguritelno pravo. Vtoro preraboteno i dopalнено izdanie*. Varna: Nauka i ikonomika.
- Andreeva, A., Yolova, G. (2019). The Challenges of the Fourth Industrial Revolution Faced by the Labour Market: European and National Processes and Trends. In *Mezhdunarodni klasterni politiki: Balgaro-kitayski forum: Sbornik s dokladi ot mezhdunarodna konferentsia*. Varna: Nauka i ikonomika.
- Andreeva, A., Yolova, G., & Dimitrova, D. (2019). Artificial intellect: Regulatory Framework and Challenges Facing the Labour Market. In *ompSysTech '19: 20-th International Conference on Computer Systems and Technologies, 21 - 22 June 2019, University of Ruse: Proceeding*. New York: ACM Digital Library.
- Andreeva, A., Yolova, G., & Dimitrova, D. (2020). Computer Technology and Ehealth. Trends and Regulatory Framework. *Economics and Law, Blagoevgrad: South-West Univ. Neofit Rilski Publ. House*, 2, 43–48.
- Andresen, L., Boud, D., & Cohen, R. (2001). Experience-Based Learning. In G. Foley (Ed.), *Understanding Adult Education and Training* (2nd ed., pp. 225–239). Allen & Unwin.
- Angusheva, V. (1987). Pravna uredba na otgovornostta na predpriyatieto za trudova zlopoluka i profesionalno zabolyavane v novia Kodeks na truda. *Pravna misal*, 33, 53-68.
- Arthur, W. B. (2013). *Complexity economics: a different framework for economic thought*. SFI Working Paper: 2013-04-012, Santa Fe Institute.
- Arthur, W. B., Durlauf, S. N., & Lane, D. A. (1997). Introduction. In W. B. Arthur, S. N. Durlauf, & D. A. Lane (Eds.), *The Economy as an Evolving Complex System II* (pp. 1–14). Addison-Wesley.
- ASA. (2019). *The causes of fraud in the financial crisis of 2007 to 2009: evidence from the mortgage-backed securities industry*. ASA. Retrieved October 2019, from <https://www.asanet.org/causes-fraud-financial-crisis-2007-2009-evidence-mortgage-backed-securities-industry>

## Compilation of References

- Association of College & Research Libraries. (2000). *Information Literacy Competency Standards for Higher Education*. Retrieved from: <https://alair.ala.org/handle/11213/7668>
- Australian Ethical Principles for Artificial Intelligence. (2019). *The Australian Government - Department of Industry, Innovation and Science*. Retrieved from: <https://consult.industry.gov.au/strategic-policy/artificial-intelligence-ethics-framework/>
- Balabanov, B. (1988). *Novi momenti v imushtestvenata otgovornost mezhdu predpriyatieto i rabotnika po Kodeksa na truda*. DP, 6.
- Balabanov, B. (1990). *Osnovaniето na iska pri smart ot trudova zlopoluka (chl.200 KT i chl.49 ZZD)*. DP,6.
- Ball Hinkelmann, K. (2016). *Modelling in Enterprise Architecture*. University of applied sciences northwest Switzerland. Business School.
- Ball, R. (1968). *The Project Gutenberg eBook of a Short Account of the History of Mathematics*. Gutenberg. Dover publications.
- Baryshnikov, Y. (2012). *IT security investment and Gordon-Loeb's I/e rule*. WEIS paper.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*. Available at: <https://gephi.org/publications/gephi-bastian-feb09.pdf>
- Beer, S. (1966). *Decision and Control – The meaning of Operational Research and Management Cybernetics*. John Wiley & Sons.
- Beinhocker, E. D. (2006). *The Origin of Wealth – Evolution, Complexity, and the Radical Remaking of Economics*. Boston: Harvard Business School Press.
- Belluzzo, R. C. B. (2014). O conhecimento, as redes e a competência em informação (CoInfo) na sociedade contemporânea: Uma proposta de articulação conceitual [Knowledge, network and information literacy (IL) in current society: a conceptual discussion]. *Perspectivas em Gestão & Conhecimento, João Pessoa*, 4, 48–63.
- Berger, J., & Rose, J. (2015). Nine Challenges for e-Government Action Researchers. *International Journal of Electronic Government Research*, 11(3), 57–75. doi:10.4018/IJEGR.2015070104
- Berge, Z., & Verneil, M. (2002). The increasing scope of training and development competency. *Benchmarking*, 9(1), 43–61. doi:10.1108/14635770210418579
- Biggs, J., & Tang, C. (2011). *Teaching For Quality Learning At University* (4th ed.). New York: McGraw Hill Society for Research into Higher Education.
- Bissell, K., Lasalle, R., & Dal Cin, P. (2020). *Innovate For Cyber Resilience*. Accenture.com. Available at: [https://www.accenture.com/\\_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf](https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf)



- Blagoycheva, H. (2020). The digitization as a stimulus for corporate social responsibility. In CSR and Socially Responsible Investing Strategies in Transitioning and Emerging Economies. IGI Global. doi:10.4018/978-1-7998-2193-9.ch003
- Blagoycheva, H., Andreeva, A., & Yolova, G. (2019). Obligation and Responsibility of Employers to Provide Health and Safety at Work – Principles, Current Regulation and Prospects. *Economic Studies*, 28(2), 115–137.
- Boguslaw, R. (1968). *New Utopians: Study of System Design and Social Change*. Spectrum Books.
- Bonsiepe, G. (2009). *Entwurfskultur und Gesellschaft: Gestaltung zwischen Zentrum und Peripherie* (1st ed.). Birkhaeuser. doi:10.1007/978-3-0346-0389-8
- Borchers, D. (2018). Das Cybersyn-Projekt Wie Chile einst die Zukunft der Planwirtschaft entwarf. *c't Retro 2018*, 77. Retrieved from <https://www.heise.de>: <https://www.heise.de/select/ct/2018/27/1541215368236612>
- Bordeleau, A. F., Mosconi, E., & Santa-Eulalia, L. A. (2018). Business Intelligence in Industry 4.0: State of the art and research opportunities. *Proceedings of the 51st Hawaii International Conference on System Sciences*. Retrieved from: <http://hdl.handle.net/10125/50383>
- Bostrom, N. (2016). *Artificial intelligence: 'we're like children playing with a bomb'*. Retrieved from: <https://www.theguardian.com/technology/2016/jun/12/nick-bostrom-artificial-intelligence-machine>
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x
- Brown, J. (2016). *Wafic Said: businessman, philanthropist and political fixer*. Financial Times. Retrieved May 15, 2017, from <https://www.ft.com/content/a3cb764a-ecf1-11e5-bb79-2303682345c8>
- Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a time of Brilliant Technologies*. Norton & Company.
- Buffomante, T. (2020). *All Hands On Deck: Key Cyber Security Considerations For 2020*. KPMG. Available at: <https://home.kpmg/xx/en/home/insights/2020/03/key-cyber-security-considerations-for-2020.html>
- Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes From the AI Frontier: Modeling the Impact of AI on the World Economy*. McKinsey Global Institute. Available at: <https://mck.co/3alkDZT>
- Buil-Gil, D., Miró-Llinares, F., Moneva, A., Kemp, S., & Díaz-Castaño, N. (2020). Cybercrime and shifts in opportunities during COVID-19: A preliminary analysis in the UK. *European Societies*, 1–13. doi:10.1080/14616696.2020.1804973

## Compilation of References

Busenitz, L. (2014). *Entrepreneurial Risk and Strategic Decision Making, It's a Matter of Perspective*. SAGE Journals.

Byalata kniga za II - Evropa v tarsene na visoki postizhenia i atmosfera na doverie (Bryuksel, 19.2.2020 COM (2020) 65 final)

Castelli, C., Gabriel, B., Yates, J., & Booth, P. (2020). *Strengthening Digital Society Against Cyber Shocks*. PwC. Available at: <https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey/strengthening-digital-society-against-cyber-shocks.html>

Catlin, T., Lorenz, J.-T., Nandan, J., Sharma, S., & Waschto, A. (2018). *Insurance beyond digital: The rise of ecosystems and platforms. Report*. McKinsey & Co.

Cearley, D., Walker, M., & Burke, B. (2016). *Top 10 Strategic Technology Trends for 2017*. Academic Press.

Chandrasekhar, A. (2018). *Trial of LTTE Financiers Begins in Switzerland-The 13 on trial face charges of fraud, false documentation, money laundering and extortion*. WIRE.

Chaplin, M., & Creasey, J. (2011). *The 2011 Standard of Good Practice for Information Security*. Information Security Forum.

Christian, B., & Griffiths, T. (2016). *Algorithms to Live by - The Computer Science of Human Decisions*. Henry Holt and Company LLC.

Chu, F., & Xie, X. (1997). *Deadlock analysis of Petri nets using siphons and mathematical programming*. *IEEE Transactions on Robotics and Automation*, 13(6), 793 - 804.

Clancey, R. (2017). *Here Lies Project Cybersyn: Salvador Allende and Stafford Beer's Cybernetic System of Coordination for Chile's Economy (1971-1973)*. Strata.

Clarke, R. (1991). Information technology and dataveillance. In C. Dunlop & R. Kling (Eds.), *Controversies in Computing*. Academic Press.

Clarke, R. (2019). Principles and processes for responsible AI. *Computer Law & Security Review*, 35(4), 410–422. doi:10.1016/j.clsr.2019.04.007

Coleman, F. (2019). *A human algorithm: how artificial intelligence is redefining who we are*. Counterpoint.

Contu, R., Canales, C., & Pingree, L. (2014). *Forecast: Information Security*. Worldwide, 2012-2018, 2Q14 Update. Gartner report, Gartner, Inc.

Cornevin, C. (2020). *La police démantèle un vaste système de blanchiment de fraude fiscale... Le Figaro*. Retrieved January 2020, from <https://www.lefigaro.fr/actualite-france/la-police-demantele-un-vaste-systeme-de-blanchiment-de-fraude-fiscale-20200110>

Daellenbach, H., McNickle, D., & Dye, Sh. (2012). *Management Science - Decision making through systems thinking* (2nd ed.). Palgrave Macmillian.

- Darling, J. R. (1999). Organizational excellence and leadership strategies: Principles followed by top multinational executives. *Leadership and Organization Development Journal*, 20(6), 309–321. doi:10.1108/01437739910292625
- Davenport, T. H. (2018). *The AI advantage: how to put the artificial intelligence revolution to work*. MIT Press. doi:10.7551/mitpress/11781.001.0001
- de Bruijn, H., & Janssen, M. (2017). Building cybersecurity awareness: The need for evidence-based framing strategies. *Government Information Quarterly*, 34(1), 1–7. doi:10.1016/j.giq.2017.02.007
- Della Croce, F., & T'kindt, V. (2002). A Recovering Beam Search algorithm for the one-machine dynamic total completion time scheduling problem. *The Journal of the Operational Research Society*, 53(11), 1275–1280. doi:10.1057/palgrave.jors.2601389
- Delponte, L. (2018). European Artificial Intelligence (AI) leadership, the path for an integrated vision. Study requested by the ITRE committee, Policy Department for Economic, Scientific and Quality of Life Policies, Directorate-General for Internal Policies, PE 626.074- September 2018.
- Demasson, A., Partridge, H., & Bruce, C. (2016). Information literacy and the serious leisure participant: Variation in the experience of using information to learn. *Information Research*, 21(2).
- Dholakia, N., Zwick, D., & Denegri-Knott, J. (2010). Technology, Consumers, and Marketing Theory. In *The SAGE Handbook of Marketing Theory* (pp. 494–511). SAGE.
- Dias, J., Khanna, S., Paquette, C., Rohr, M., Seitz, B., Singla, A., Sood, R., & van Ouwerkerk, J. (2017). *Introducing The Next-Generation Operating Model*. Available at: <https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/introducing%20the%20next-generation%20operating%20model/introducing-the-next-gen-operating-model.ashx>
- Dimitrova, D., Mateeva, Z., & Dimitrova, D. (2020). *Administrativno pravo i protses*. Varna: Nauka i ikonomika.
- Dodge, R. C. Jr, Carver, C., & Ferguson, A. J. (2007). Phishing for user security awareness. *Computers & Security*, 26(1), 73–80. doi:10.1016/j.cose.2006.10.009
- Dogan, S., Çalgici, P., Arditi, D., & Gunaydin, H. (2015). *Critical success factors of partnering in the building design process*. METU JFA 2015/2. Department of Architecture, İzmir Institute of Technology.
- Dowling, S., Schukat, M., & Barrett, E. (2018). Improving adaptive honeypot functionality with efficient reinforcement learning parameters for automated malware. *Journal of Cyber Security Technology*, 2(2), 75–91. doi:10.1080/23742917.2018.1495375
- Dunn, J. (2012). *The importance of internet access in schools*. Available at <http://www.edudemic.com/2012/12/the-importance-of-internet-access-in-schools/>
- Duparc, A. (2010). *La Suisse restitue au Liban les archives du fonds Dunand*. Le Monde.

## Compilation of References

- Easterbrook, S., Singer, J., Storey, M., & Damian, D. (2008). *Guide to Advanced Empirical Software Engineering-Selecting Empirical Methods for Software Engineering Research*. Springer.
- El Hashem, B. (1990). *It was Kissinger who destroyed the nation of Lebanon*. EIR Feature.
- Eminağaoğlu, M., Uçar, E., & Eren, S. (2009). The positive outcomes of information security awareness training in companies – A case study. *Information Security Technical Report*, 14(4), 223–229. doi:10.1016/j.istr.2010.05.002
- Erboz, G. (2017). How To Define Industry 4.0: Main Pillars Of Industry 4.0. *Conference: 7th International Conference on Management (ICoM 2017)*.
- Espejo, R. (2017). Cybernetic Argument for Democratic Governance: Cybersyn and Cyberfolk. In L. C. Werner (Ed.), *Con-Versations Vol.1 cybernetics: state of the art* (pp. 34-57). Berlin: Universitätsverlag der TU Berlin.
- Ethical Guideliness for Trustworthy Artificial Intelligence. (2019). *European Comission. High-Level Expert Group on Artificial Intelligence (AI HLEG)*. Retrieved from: <https://ec.europa.eu/futurium/en/ai-alliance-consultation>
- European commission, (2004). Commission of the European Communities Legal barriers in e-business: The results of an open consultation of enterprises. Brussels, 26.4.2004 SEC(2004) 498. European commission.
- EvansP. C. AnnunziataM. (2012). Retrieved from [https://www.ge.com/docs/chapters/Industrial\\_Internet.pdf](https://www.ge.com/docs/chapters/Industrial_Internet.pdf)
- Feldforth, O. (2019). *Arbeitszeit klar erfassen - aber wie?* Retrieved from <https://www.tagesschau.de/>: <https://www.tagesschau.de/wirtschaft/eugh-arbeitszeiten-107.html>
- Felfel, H., Ayadi, O., & Masmoudi, F. (2017). Pareto Optimal Solution Selection for a Multi-Site Supply Chain Planning Problem Using the VIKOR and TOPSIS Methods. *International Journal of Service Science, Management, Engineering, and Technology*. Doi:10.4018/IJSSMET.2017070102
- FireEye. (2020). *Security Effectiveness 2020: Deep Dive Into Cyber Security Reality*. <https://content.fireeye.com/security-effectiveness/rpt-security-effectiveness-2020-deep-dive-into-cyber-reality>
- Fischer, T. (2019). Kybernetik. In T. Schoeler, S. Hoeltgen, & J. F. Maibaum (Eds.), *Medientechnisches Wissen* (pp. 275–301). De Gruyter.
- Fishbein, M., & Ajzen, I. (2011). *Predicting and Changing Behavior: The reasoned Action Approach*. Psychology Press. doi:10.4324/9780203838020
- Fitsanakis, J. (2016). *Switzerland made secret deal with PLO in the 1970s, new book alleges*. Academic Press.
- Flatow, I. (2008). *Web privacy concerns prompt Facebook changes*. In *On Science Friday*. NPR ScienceFriday Inc.

- Floridi, L. (2004). Open problems in the philosophy of information. *Metaphilosophy*, 35(4), 2004. doi:10.1111/j.1467-9973.2004.00336.x
- Frey, C. B., & Osborne, M. A. (2013). *The Future Employment: How susceptible are jobs to computerisation?* Retrieved from <https://www.oxfordmartin.ox.ac.uk>: [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf)
- Fuduric, M., & Mandelli, A. (2014). Communicating social media policies: Evaluation of current practices. *Journal of Communication Management (London)*, 18(2), 158–175. doi:10.1108/JCOM-06-2012-0045
- Gartner. (2013). *Scenario Toolkit: Using EA to Support Business Transformation*. Gartner Inc.
- Gartner. (2016). *Gartner's 2016 Hype Cycle for ICT in India Reveals the Technologies that are Most Relevant to Digital Business in India Analysts to Explore Key Technologies and Trends*. Gartner Symposium/ITxpo 2016, Goa, India.
- Gatlan, S. (2020). Coronavirus Phishing Attacks Are Actively Targeting the US. *Bleeping Computer*. <https://www.bleepingcomputer.com/news/security/coronavirus-phishing-attacks-are-actively-targeting-the-us/>
- Giachetti, R. (2012). *A Flexible Approach to Realize an Enterprise Architecture*. Department of Systems Engineering, Naval Postgraduate School, Monterey, CA USA. Presented: New Challenges in Systems Engineering and Architecting. Conference on Systems Engineering Research (CSER). St. Louis, MO. 10.1016/j.procs.2012.01.031
- Gintis, H. (2006). *The economy as a complex adaptive system. A Review of Eric D. Beinhocker The Origins of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*.
- Glanville, R. (2012). Radical constructivism = second-order Cybernetics. *Cybernetics & Human Knowing*, 4(4), 27–42.
- Goikoetxea, A. (2004). A mathematical framework for enterprise architecture representation and design. *International Journal of Information Technology & Decision Making*, 03(01), 5–32. doi:10.1142/S0219622004000623
- Goleman, D. (1995). *Emotional intelligence*. Bantam Books.
- Goleva, P. (2011). *Deliktno pravo*. Feneya.
- Grafstein, A. (2017). Information Literacy and Critical Thinking: Context and Practice. Pathways into Information Literacy and Communities of Practice: Teaching Approaches and Case Studies, 3-28.
- Greethorst, D. (2009). *Using the Open Group's Architecture Framework as a pragmatic approach to architecture*. *Jaarbeurs, Utrecht. KIVI NIRIA, afd. Informatica*.
- Griffiths, T., & Christian, B. (2016). *Algorithms to Live By: The Computer Science of Human Decisions*. Henry Holt and Co.

## Compilation of References

- Gunasekare, U. (2015). *Mixed Research Method as the Third Research Paradigm: A Literature Review*. University of Kelaniya.
- Guzdial, M. (2015). *Using learning sciences to inform cyber security education*. Georgia Tech College of Computing. <https://computing.wordpress.com/2015/05/18/using-learning-sciences-to-inform-cyber-security-education/>
- Hakak, S., Khan, W. Z., Imran, M., Choo, K. K. R., & Shoaib, M. (2020). Have You Been a Victim of COVID-19-Related Cyber Incidents? Survey, Taxonomy, and Mitigation Strategies. *IEEE Access: Practical Innovations, Open Solutions*, 8, 124134–124144. doi:10.1109/ACCESS.2020.3006172
- Hamel, G. (2002). *Leading The Revolution*. Plume.
- Harari, Y. N. (2017). *Homo deus: a brief history of tomorrow*. Harper. doi:10.17104/9783406704024
- Harley, D. (2010). *Re-floating the Titanic: dealing with social engineering attacks*. <https://smallbluegreenblog.files.wordpress.com/2010/04/eicar98.pdf>
- Härtig Attorneys at Law. (2013). *Legal aspects of social media*. Available at: [https://www.haerting.de/sites/default/files/downloads/handout\\_legal\\_aspects\\_of\\_social\\_media\\_2013.pdf](https://www.haerting.de/sites/default/files/downloads/handout_legal_aspects_of_social_media_2013.pdf)
- Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. (2013). *The Atlas of the Economic Complexity: Mapping Paths to Prosperity*. Massachusetts Institute of Technology and Center for International Development, Harvard University.
- Herzog, P. (2003). *OSSTMM 2.1 Open-Source Security Testing Methodology Manual*. Institute for Security and Open Methodologies.
- Hessler, N. (2014). Die Halle 54 bei Volkswagen und die Grenzen der Automatisierung. *Zeithistorische Forschungen/Studies in Contemporary History*, 11, 56-76. Retrieved from <https://zeithistorische-forschungen.de/1-2014/id%3D4996>
- Hinduja, S., & Patchin, J. W. (2014). *Cyberbullying*. Cyberbullying Research Center. Retrieved September, 7, 2015, retrieved from <https://cyberbullying.org/Cyberbullying-Identification-Prevention-Response.pdf>
- Ho, W., Xu, X., & Dey, P. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. Operations and Information Management Group, Aston Business School, Aston University.
- Hobbs, K. (2020). Socially distancing from COVID-19 robocall scams. *Consumer Information*. <https://www.consumer.ftc.gov/blog/2020/03/socially-distancing-covid-19-robocall-scams>
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture*, 2(1), 2307–2319. doi:10.9707/2307-0919.1014
- Human intelligence definition. (n.d.). Retrieved from [https://en.wikipedia.org/wiki/Human\\_intelligence](https://en.wikipedia.org/wiki/Human_intelligence)

- Imgraben, J., Engelbrecht, A., & Choo, K. (2014). Always connected, but are smart mobile users getting more security savvy? A survey of smart mobile device users'. *Behaviour & Information Technology*, 33(12), 1347–1360. doi:10.1080/0144929X.2014.934286
- Introna, L. D. (2007). Maintaining the reversibility of folding: Making the ethics (politics) of information technology visible. *Ethics and Information Technology*, 9(1), 11–25. doi:10.1007/10676-006-9133-z
- Irwin, L. (2020). *List of Data Breaches and Cyber Attacks in February 2020–623 million Records Breached*. IT Governance.
- Izkustven intelekt za Evropa COM/2018/237 final Saobshtenie na komisiyata do evropeyskia parlament, saveta, ikonomicheskia i sotsialen komitet i komiteta na regionite. <https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=CELEX%3A52018DC0237>
- Izzo, S. (2019). Karl-Heinz Hoffmann's Secret History Links Neo-Nazis With Palestinian Terror - Tablet Magazine. *Tablet (Brooklyn, N.Y.)*. Retrieved October 2019, from <https://www.tabletmag.com/jewish-arts-and-culture/culture-news/286220/karl-heinz-hoffmann-far-right> 1/11
- Jara, A. J., Parra, M. C., & Skarmeta, A. F. (2012). Marketing 4.0: A New Value Added to the Marketing through the Internet of Things. *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on*, 852-857.
- Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality & Quantity*, 41(1), 37–54. Retrieved August 10, 2018, from <https://link.springer.com/article/10.1007/s11135-005-5427-1>
- Jochum, G. (2017). *Plus Ultra« oder die Erfindung der Moderne: Zur neuzeitlichen Entgrenzung ...* Bielefeldt: transcript.
- Johnson, R., & Onwuegbuzie, A. (2004). *Mixed Methods Research: A Research Paradigm Whose Time Has Come*. Sage Journals.
- Johnson, S. (1983). Francois Genoud: Terrorist controller for Swiss banks. *Executive Intelligence Review*.
- Jonkers, H., Band, I., & Quartel, D. (2012a). *ArchiSurance Case Study*. The Open Group.
- Joseph, C. (2014). *Types of eCommerce Business Models*. Retrieved September 17, 2019, from <https://smallbusiness.chron.com/types-ecommerce-business-models-2447.html>
- Jurse, M., & Mulej, M. (2011). The complexities of business school alignment with the emerging globalisation of business education. *Kybernetes*, 40(9/10), 1440–1458. doi:10.1108/03684921111169477
- Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review*. Available at: <<https://sloanreview.mit.edu/digital2015>>

## Compilation of References

- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59–68. doi:10.1016/j.bushor.2009.09.003
- Kaufmann, S. (2016). *Digitalisierung, Klassenkampf, Revolution*. Retrieved from <https://www.rosalux.de>: [https://www.rosalux.de/fileadmin/images/publikationen/Analysen/Analysen33\\_Digitalisierung.pdf](https://www.rosalux.de/fileadmin/images/publikationen/Analysen/Analysen33_Digitalisierung.pdf)
- Khan, W. Z., Khan, M. K., Muhaya, F. T. B., Aalsalem, M. Y., & Chao, H. C. (2015). A comprehensive study of email spam botnet detection. *IEEE Communications Surveys and Tutorials*, 17(4), 2271–2295. doi:10.1109/COMST.2015.2459015
- Kietzmann, J., Hermkens, K., McCarthy, I., & Silvestre, B. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), 241–251. doi:10.1016/j.bushor.2011.01.005
- 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
- Klein, G. (2016). Trying to make rational decisions while employing intuitive reasoning: A Look at the due-diligence process using the dual-system reasoning model. *International Journal of Entrepreneurship and Innovation Management*, 20(3/4), 214–234. doi:10.1504/IJEIM.2016.077962
- Klein, G., & Shtudiner, Z. (2016). Trust in others: Does it affect investment decisions? *Quality & Quantity*, 50(5), 1949–1967. doi:10.1007/11135-015-0245-6
- Knight, W. (2017). *Boston may be famous for bad drivers, but it's the testing ground for a smarter self-driving car*. MIT Technology Review.
- Koering, D. (2019). Conscious City Laboratory - Explorations in the history of computation, cybernetics, and architecture: Foresight for artificial intelligence and human participation within cities. Universitätsverlag der TU Berlin. DOI: 10.14279/depositonce-8466
- Kordon, A. (2012, October). *Applying data mining in raw materials forecasting*. Paper presented at the SAS Analytics 2012 Conference, Las Vegas, NV.
- Kordon, A. K. (2020). *Applying data science: how to create value with artificial intelligence*. Springer.
- Kozhuharov, A. (1958). *Obligatsionno pravo. Obshto uchenie za obligatsionnoto otmoshenie*. Sofia: Nauka i izkustvo.
- Kraisig, A., Rosélia, A., Welter, F., Haugg, I., Cargnin, R., Roos-Frantz, F., Sawicki, S., & Frantz, R. (2016). Mathematical Model for Simulating an Application Integration Solution in the Academic Context of Unijui University. *Procedia Computer Science*, 100, 407–413. doi:10.1016/j.procs.2016.09.176



- Kumaraguru, P., Acquisti, A., Rhee, Y., & Nunge, E. (2007). Protecting people from phishing: The design and evaluation of an embedded training email system. *Conference on Human Factors in Computing Systems – Proceedings*, 905-914. 10.1145/1240624.1240760
- Kuperman, G. J., Reichley, R. M., & Bailey, T. C. (2006). Using Commercial Knowledge Bases for Clinical Decision Support: Opportunities, Hurdles, and Recommendations. *Journal of the American Medical Informatics Association*, 13(4), 369–371. doi:10.1197/jamia.M2055 PMID:16622160
- Lamarre, T. (2012). Humans and Machines. *Inflexions*, 5, 29–67.
- Lau, J. (2007). *Guidelines on Information Literacy for Lifelong Learning*. The Hague: IFLA. Retrieved from: <https://www.ifla.org/files/assets/information-literacy/publications/ifla-guidelines-en.pdf>
- 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
- Le Monde. (2019). *Le Prix Nobel d'économie Angus Deaton: « Quand l'Etat produit une élite prédatrice »*. Le Monde. Retrieved November 14, 2019, from [https://www.lemonde.fr/idees/article/2019/12/27/angus-deaton-quand-l-etat-produit-une-elite-predatrice\\_6024205\\_3232.html](https://www.lemonde.fr/idees/article/2019/12/27/angus-deaton-quand-l-etat-produit-une-elite-predatrice_6024205_3232.html)
- Le News. (2015). *Swiss People's Party (UDC) leaders found guilty of racism*. Le News. Retrieved September 2019, from <https://lenews.ch/2015/04/30/two-swiss-peoples-party-udc-leaders-found-guilty-of-racism/>
- Le News. (2017). *Racism sentence upheld against former Swiss People's Party secretary general*. Le News. Retrieved September 2019, from <https://lenews.ch/2017/04/13/racism-sentence-upheld-against-former-swiss-peoples-party-secretary-general/>
- Leitch, R. & Day, C. (2000). *Action research and reflective practice: towards a holistic view*. Taylor & Francis.
- Letho, M. (2015), Cyber Security Competencies - Cyber Security Education and Research in Finnish Universities. *14th European Conference on Cyber Warfare and Security (ECCWS)*, 179-188.
- Levin, S. (2019). *Google scraps AI ethics council after backlash: Back to the drawing board*. <https://www.theguardian.com/technology/2019/apr/04/google-ai-ethics-council-backlash>
- List of cognitive biases. (n.d.). Retrieved from [https://en.wikipedia.org/wiki/List\\_of\\_cognitive\\_biases](https://en.wikipedia.org/wiki/List_of_cognitive_biases)
- Lloyd, A. (2007). Recasting information literacy as sociocultural practice: Implications for library and information science researchers. *Information Research*, 12(4).
- Lockwood, R. (2018). *Introduction The Relational Data Model*. Retrieved January 29, 2019, from <http://www.jakobsens.dk/Nekrologer.htm>

## Compilation of References

- Loginovskiy, O. V., Dranko, O. I., & Holloy, A. V. (2018). *Mathematical Models for Decision-Making on Strategic Management of Industrial Enterprise in Conditions of Instability*. Conference: Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications (EMIT 2018). Leipzig, Germany.
- Luckner, A. (2008). *Heidegger und das Denken der Technik*. Bielefeldt: Transcript.
- Lukanović, L. (2017). *Računalniška kriminaliteta in varstvo osebnih podatkov: diplomska naloga*. Available at: [http://www.ediplome.fm-kp.si/Lukanovic\\_Lea\\_20171017.pdf](http://www.ediplome.fm-kp.si/Lukanovic_Lea_20171017.pdf)
- Mangold, W., & Faulds, D. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, 52(4), 357–365. doi:10.1016/j.bushor.2009.03.002
- Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The Internet Of Things: Mapping The Value Beyond The Hype*. McKinsey Global Institute. Available at: <https://www.mckinsey.com/business-functions/business-technology/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world>
- Marcus, G., & Davis, E. (2019). *Rebooting AI: building artificial intelligence we can trust*. Pantheon Books.
- Markides, C. C. (2015). Research on Business Models: Challenges and Opportunities. *Advances in Strategic Management*, 33, 133–147. doi:10.1108/S0742-332220150000033004
- McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, (October), 2012. PMID:23074865
- Mccarthy, J. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence, 1955*. Retrieved from: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
- McCrohan, K., Engel, K., & Harvey, J. (2010). Influence of awareness and training on cyber security. *Journal of Internet Commerce*, 9(1), 23–41. doi:10.1080/15332861.2010.487415
- McDaniel, E. (2013, July). Securing the information and communications technology global supply chain from exploitation: Developing a strategy for education, training, and awareness. In *Proceedings of the Informing Science and Information Technology Education Conference* (pp. 313-324). Informing Science Institute. doi:10.28945/1813
- Medina, E. (2011). *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. MIT Press. doi:10.7551/mitpress/8417.001.0001
- 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. doi:10.1145/1101908.1101940
- Menon, S. (2020, April 19). Coronavirus: Herbal remedies in India and other claims fact-checked. *BBC News*. <https://www.bbc.com/news/world-asia-india-51910099>

- Miller, A. G. (1955). The Magical Number Seven, Plus or Minus Two Some Limits on Our Capacity for Processing Information. *Psychological Review*, 101. Retrieved from Psychological Review: <https://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf>
- Mitchell, M. (2019). *Artificial intelligence: a guide to thinking human*. Farrar, Strauss and Giroux.
- Montgomery, M. (1999). Complexity Theory: An Austrian Perspective. In D. Colander (Ed.), *Complexity Theory and the History of Economic Thought*. Routledge Press. <http://www.rasmusen.org/xpacioli/workpaps/99.06.Complexity.PDF>
- Moore, J. (2014). *Java programming with lambda expressions-A mathematical example demonstrates the power of lambdas in Java 8*. Retrieved March 10, 2018, from, <https://www.javaworld.com/article/2092260/java-se/java-programming-with-lambda-expressions.html>
- Morozov, E. (2014). *The Planning Machine: Project Cybersyn and the origins of the Big Data nation*. The New Yorker.
- Mrachkov, V. (2013). *Imushchestvena otgovornost na rabotodatelya*. Sibi.
- Myers, B., Pane, J., & Ko, A. (2004). *Natural programming languages and environments*. ACM New York. doi:10.1145/1015864.1015888
- Nascimento Marques, M. R., Jr. (2018). Embedded Agent based on Cyber Physical Systems: Architecture, Hardware Definition and Application in Industry 4.0 Context. In *15th International Conference on Informatics in Control, Automation and Robotics* (pp. 584-591). Retrieved from Center of Computational Sciences, Federal University of Rio Grande, Rio Grande, Brazil: [www.researchgate.ent](http://www.researchgate.ent)
- Neumann, G. (2002). Programming Languages in Artificial Intelligence. In *Encyclopaedia of Information Systems*. Academic Press.
- Neumann, B., & Moller, R. (2008). On scene interpretation with description logics. *Image and Vision Computing*, 26(1), 82–101. doi:10.1016/j.imavis.2007.08.013
- 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
- Nilda Tri, P., & Yusof, S. M. (2009). Critical Success Factors for Implementing Quality Engineering Tools and Techniques in Malaysian's and Indonesian's Automotive Industries: An Exploratory Study. In *Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2009*. MECS.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review (Seattle, Wash.)*, 79, 101–139.

## Compilation of References

- NIST Roadmap for Improving Critical Infrastructure Cybersecurity. (2014). <https://www.nist.gov/cyberframework/upload/roadmap-021214.pdf>
- North, N. (2010). *Behaviour-Driven Development Writing software that matters*. DRW publications.
- Obama, B. (2009). *National information literacy awareness month*. Retrieved from: <https://www.govinfo.gov/app/content/pkg/STATUTE-123/pdf/STATUTE-123-Pg3711.pdf>
- OECD. (2011). *Global Forum on Transparency and Exchange of Information for Tax Purposes Peer Review: Switzerland 2011, Phase I*. OECD Publishing.
- OECD. (2014). *HEInnovate: Introduction to HEInnovate and its seven dimensions*. Available at: <https://www.oecd.org/cfe/leed/HEInnovate-Introduction%20.pdf>
- OECD. (2018). *Country Case Study 1: Lebanon. MENA-OECD Economic Resilience Task Force Resilience In Fragile Situations. 4-5 December 2018. Islamic Development Bank*. OECD.
- Olley, D. (2020). *AI Report | Research Intelligence | Elsevier*. <https://www.elsevier.com/research-intelligence/resource-library/ai-report>
- Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the 'IT' in IT research – a call to theorizing the IT artifact. *Information Systems Research*, 12(2), 121–134. doi:10.1287/isre.12.2.121.9700
- Otonicar, S. L. C., Valentim, M. L. P., & Feres, G. G. (2015). Competência em informação e os contextos educacional, tecnológico, político e organizacional [Information literacy and the educational, technological, political and organizational contexts]. *Revista Ibero-americana de Ciência da Informação, Brasília*, 9(1), 24–142.
- Panetta, K. (2018). *Ecosystems Drive Digital Growth*. <https://www.gartner.com/smarterwithgartner/ecosystems-drive-digital-growth/>
- Paravicini, G. (2018). *Millions flow from Gaddafi's 'frozen funds' to unknown beneficiaries*. Politico. <https://www.politico.eu/article/muammar-gaddafi-frozen-funds-belgium-unknown-beneficiaries/>
- Parisi, G. (1999). Complex systems: A physicist's viewpoint. *Physica A*, 263(1), 557–564. doi:10.1016/S0378-4371(98)00524-X
- Pawlowski, S., & Yoonhyuk, J. (2015). Social representations of cyber security by university students and implications for instructional design. *Journal of Information Systems Education*, 26(4), 281–294.
- Peterson, S. (2011). *Why it Worked: Critical Success Factors of a Financial Reform Project in Africa*. Faculty Research Working Paper Series. Harvard Kennedy School.
- Philp, W., & Martin, C. (2009). A philosophical approach to time in military knowledge management. *Journal of Knowledge Management*, 13(1), 171–183. doi:10.1108/13673270910931242
- Pias, C. (2007). *Defense of Cybernetics. A Reminiscence*. WEB.

- Pierson, J., & Heyman, R. (2011). Social media and cookies: Challenges for online privacy. *Info*, 13(6), 30–42. doi:10.1108/14636691111174243
- Polderman, J., & Willems, J. (1998). Introduction to Mathematical Systems Theory: A Behavioral Approach. Springer Verlag. Germany. doi:10.1007/978-1-4757-2953-5
- Prakken, H. (2017). On the problem of making autonomous vehicles conform to traffic law. *Artificial Intelligence and Law*, 25(3), 341–363. doi:10.1007/10506-017-9210-0
- Predloženie za rezolyutsia na evropeyskia parlament sadarzhashto preporaki kam Komisiyata otnosno grazhdanskopravni normi za robotikata (2015/2103(INL)). [https://www.europarl.europa.eu/doceo/document/A-8-2017-0005\\_BG.html#title1](https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_BG.html#title1)
- Rader, C., Subhan, Z., Lanier, C., Brooksbank, R., Yankah, S., & Spears, K. (2014). CyberRx. *International Journal of Pharmaceutical and Healthcare Marketing*, 8(2), 193–225. doi:10.1108/IJPHM-05-2013-0027
- Ravanetti, A. (2016). *Switzerland Bank on Fintech with Lighter Regulations*. Crowd Valey. Retrieved September 2019, from <https://news.crowdvalley.com/news/switzerland-bank-on-fintech-with-lighter-regulations>
- Rek, M., & Milanovski, B.K. (2017). *Slovenija, Ljubljana: Fakulteta za medije* [izdelava]. Slovenija, Ljubljana: Univerza v Ljubljani, Arhiv družboslovnih podatkov [distribucija], IDNo: MPSS16.
- Renaud, K., & Weir, G. R. (2016). Cybersecurity and the Unbearability of uncertainty. *2016 Cybersecurity and Cyberforensics Conference (CCC)*. 10.1109/CCC.2016.29
- Reuters. (2019a). *Swiss group files criminal complaint against Credit Suisse over Mozambique loans*. Reuters. <https://www.reuters.com/article/us-mozambique-creditsuisse/swiss-group-files-criminal-complaint-against-credit-suisse-over-mozambique-loans-idUSKCN1S5174>
- Rickman, R. (1999). *Swiss Banks and Jewish Souls by Gregg J. Rickman*. Central European History. JSTOR.
- Rinat & Vardan. (2019a). Math model of neuron and nervous system research, based on AI constructor creating virtual neural circuits: Theoretical and Methodological Aspects. In V. Mkrttchian, E. Aleshina, & L. Gamidullaeva (Eds.), *Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms* (pp. 320–344). IGI Global. doi:10.4018/978-1-7998-1581-5.ch015
- Rinat & Vardan. (2019b). Brain machine interface – for Avatar Control & Estimation in Educational purposes Based on Neural AI plugs: Theoretical and Methodological Aspects. In V. Mkrttchian, E. Aleshina, & L. Gamidullaeva (Eds.), *Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms* (pp. 345–360). IGI Global. doi:10.4018/978-1-7998-1581-5.ch016
- Rinat, G. (2020). Brain machine interface: The accurate interpretation of neurotransmitters' signals targeting the muscles. *International Journal of Applied Research in Bioinformatics*, 0102. Advance online publication. doi:10.4018/IJARB.2020

## Compilation of References

- RSAC. (2019). *The Future of Companies and Cybersecurity Spending*. RSA Conference. Available at: <https://www.rsaconference.com/industry-topics/blog/the-future-of-companies-and-cybersecurity-spending#:~:text=In%202019%2C%20worldwide%20spending%20on,to%20reach%20over%20%24124%20billion.&text=Spending%20on%20security%20services%20has%20reached%20%2464.2%20million%20in%202019>
- Russel, S. (2019). *Human compatible: artificial intelligence and the problem of control*. Viking.
- Sankaralingam, K., Ferris, M., Nowatzki, T., Estan, C., Wood, D., & Vaish, N. (2013). *Optimization and Mathematical Modeling in Computer Architecture*. Morgan & Claypool Publishers.
- Scherer, R. J., & Schapke, S. E. (2011, October). A distributed multi-model-based Management Information System for simulation and decision making on construction projects. *Advanced Engineering Informatics*, 25(4), 582–599. doi:10.1016/j.aei.2011.08.007
- Schwaab, K. (2017). *The Fourth Industrial Revolution*. Crown Business.
- Schwab, K. (2016). *The fourth industrial revolution*. Crown Business.
- Schwartz, E. M. (2010, May). *Poverty reduction for profit? A critical assessment of the Bottom-of-the-Pyramid Approach and of the 'Opportunities for the Majority'-Initiative of the Inter-American Development Bank*. Retrieved from University Vienna: <https://core.ac.uk/download/pdf/11590712.pdf>
- Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2019). An evaluation framework for network security visualizations. *Computers & Security*, 84, 70–92. doi:10.1016/j.cose.2019.03.005
- Sharf, E. (2016). Information exchanges: regulatory changes to the cyber-security industry after Brexit: Making security awareness training work. *Computer Fraud & Security*, 2016(7), 9–12. doi:10.1016/S1361-3723(16)30052-5
- Shaw, R. S., Chen, C. C., Harris, A. L., & Huang, H.-J. (2009). The impact of information richness on information security awareness training effectiveness. *Computers & Education*, 52(1), 92–100. doi:10.1016/j.compedu.2008.06.011
- Shelton, C. (1999). *Quantum Leaps*. Butterworth-Heinemann.
- Shelton, C., & Darling, J. (2004). From chaos to order: Exploring new frontiers in conflict management. *Organization Development Journal*, 22(3), 22–41.
- Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J., Lyons, T., Etchemendy, J., Grosz, B., & Bauer, Z. (2018). *The AI Index 2018 Annual Report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University.
- Shtudiner, Z., Klein, G., Zwilling, M., & Kantor, J. (2019). The value of souvenirs: Endowment effect and religion. *Annals of Tourism Research*, 74, 17–32. doi:10.1016/j.annals.2018.10.003

- 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, J. (2006). Framing brand management for marketing ecosystems. *The Journal of Business Strategy*, 27(5), 50–57. doi:10.1108/02756660610692716
- Slayton, R. (2018). *Policy Series: Beyond Cyber-Threats: The Technopolitics of Vulnerability*. Retrieved from: <https://issforum.org/roundtables/policy/1-5bc-technopolitics>
- Smith, G. (2018). *The AI delusion*. Oxford University Press. doi:10.1093/oso/9780198824305.001.0001
- Smoliar, S. W., & HongJiang Zhang. (1994). Content based video indexing and retrieval. *IEEE MultiMedia*, 1(2), 62–72. doi:10.1109/93.311653
- Snowden, E. (2015). *Most Racist, Award Goes To ... Switzerland? Skating on Stilts*. Retrieved September 2019, from <https://www.skatingonstilts.com/skating-on-stilts/2015/03/and-the-edward-snowden-most-racist-award-goes-to-switzerland.html>
- Solove, D. J. (2004). *The Digital Person: Technology and Privacy in the Information Age*. University Press.
- Sredkova, K. (2011). *Trudovo pravo. Spetsialna chast. Dyal I. Individualno trudovo pravo*. Sofia: UI Sv. Kl. Ohridski.
- Srinivas, J., Das, A. K., & Kumar, N. (2019). Government regulations in cyber security: Framework, standards and recommendations. *Future Generation Computer Systems*, 92, 178–188. doi:10.1016/j.future.2018.09.063
- Stein, S., & Jacobs, J. (2020, March 16). *Cyber-attack hits U.S. health agency amid COVID-19 outbreak*. Bloomberg.com. <https://www.bloomberg.com/news/articles/2020-03-16/u-s-health-agency-suffers-cyber-attack-during-covid-19-response>
- Stempel, J. (2019). *UBS must defend against U.S. lawsuit over 'catastrophic' mortgage losses*. Yahoo Finance. Retrieved September 2019, from <https://finance.yahoo.com/news/ubs-must-defend-against-u-214743943.html>
- Stern, J. (2017). *Artificial intelligence for marketing*. John Wiley & Sons. doi:10.1002/9781119406341
- Stowasser, S. (2019). *KI verändert die Welt – auch die Arbeit*. Retrieved from <https://www.wissenschaftsjahr.de/2019/neues-aus-der-wissenschaft/das-sagt-die-wissenschaft/ki-veraendert-die-welt-auch-die-arbeit/>
- Stupples, B., Sazonov, A., & Woolley, S. (2019). *UBS Whistle-Blower Hunts Trillions Hidden in Treasure Isles*. *Bloomberg-Economics*. Bloomberg. Reviewed in November 2019 <https://www.bloomberg.com/news/articles/2019-07-26/ubs-whistle-blower-hunts-trillions-hidden-in-treasure-islands>

## Compilation of References

- Sweeting, B. (2019, April 1). Applying ethics to itself: Recursive ethical questioning in architecture and second-order cybernetics. *Kybernetes*, 48(4), 805–815. Advance online publication. doi:10.1108/K-12-2017-0471
- Syynimaa, N. (2015). *Enterprise Architecture Adoption Method for Higher Education Institutions* (Doctoral Thesis). Informatics Research Centre Henley Business School University of Reading.
- Tabansky, L. (2013). Critical Infrastructure Protection Policy: The Israeli Experience. *Journal of Information Warfare*, 12(3), 78–86.
- Tegmark, M. (2017). *Life 3.0 - Being Human in the Age of Artificial Intelligence*. Allan Lane.
- Tegmark, M. (2017). *Life 3.0 - Being human in the age of Artificial Intelligence*. Penguin Random House UK.
- Thakur, K., & Pathan, A. S. K. (1995, August 10). Cybersecurity fundamentals: A real-world perspective”. Routledge & CRC Press. <https://www.routledge.com/Cybersecurity-Fundamentals-A-Real-World-Perspective/Thakur-Pathan/p/book/9780367472504>
- The Australian Government - Department of Industry. (2019). *Innovation and Science. AI Ethics Principles*. Retrieved from: <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>
- The Open Group. (2011a). *Architecture Development Method*. The Open Group. Reviewed in February 2018, <https://pubs.opengroup.org/architecture/togaf9-doc/arch/chap05.html>
- The Open Group. (2011a). *Architecture Development Method*. The Open Group. USA. Reviewed in February 2018, <https://pubs.opengroup.org/architecture/togaf9-doc/arch/chap05.html>
- The Open Group. (2011b). *TOGAF 9.1*. The Open Group. Reviewed in August 2018, <http://www.opengroup.org/subjectareas/enterprise/togaf>
- Thomas, A. (2015). *Gartner, Innovation Insight for Microservices*. Gartner.
- Tidd, J. (2006). *From Knowledge Management to Strategic Competence* (2nd ed.). Imperial College. doi:10.1142/p439
- Tidd, J., & Bessant, J. (2009). *Managing Innovation, Integrating Technological, Market and Organizational Change* (4th ed.). Wiley.
- Tipton, W.H. (2014). Cyber security education: remove the limits. *Information Week*. <https://www.informationweek.com/government/cybersecurity/cyber-security-education-remove-the-limits/a/d-id/1306950>
- Trad, A. (2018a). *The Business Transformation and Enterprise Architecture Framework-Applied to analyse / The historically recent Rise and the 1975 Fall of the Lebanese Business Ecosystem*. IGI-Global.



- Trad, A. (2019d). *Applied Mathematical Model for Business Transformation-Assessing Risks of the Lebanese Islamic Business/Marketing Strategy and its Relationships with Counterparts/ Partners (IBM&R)*. IGI-Global.
- Trad, A. (2019e). *Applied Mathematical Model for Business Transformation Projects-The intelligent Strategic Decision Making System (iSDMS)*. Encyclopaedia. IGI-Global.
- Trad, A., & Kalpić, D. (2017b). *An Intelligent Neural Networks Micro Artefact Patterns' Based Enterprise Architecture Model*. IGI-Global.
- Trad, A., & Kalpić, D. (2017c). *A Neural Networks Portable and Agnostic Implementation TKM&F for Business Transformation Projects. The Framework*. IEEE.
- Trad, A., & Kalpić, D. (2017d). *A Neural Networks Portable and Agnostic Implementation TKM&F for Business Transformation Projects- The Basic Structure*. IEEE Conference on Computational Intelligence.
- Trad, A., & Kalpić, D. (2018a). *The Business Transformation Framework and Enterprise Architecture Framework for Managers in Business Innovation-Knowledge and Intelligence Driven Development (KIDD)*. Encyclopaedia of E-Commerce Development, Implementation, and Management. IGI-Global.
- Trad, A., & Kalpić, D. (2018b). *The Business Transformation Framework and Enterprise Architecture Framework for Managers in Business Innovation- Knowledge Management in Global Software Engineering (KMGSE)*. Encyclopaedia of E-Commerce Development, Implementation, and Management. IGI-Global.
- Trad, A., & Kalpić, D. (2019c). *Business Transformation and Enterprise Architecture-The Resources Management Research and Development Project (RMSRDP)*. Book. IGI-Global.
- Trad, A., & Kalpić, D. (2019e). *Business Transformation and Enterprise Architecture-The Holistic Project Resources Management Pattern (HPRMP)*. Encyclopaedia. IGI-Global.
- Trad, A., & Kalpić, D. (2020a). *Using Applied Mathematical Models for Business Transformation*. IGI Complete Author Book. IGI Global. doi:10.4018/978-1-7998-1009-4
- Trading Economics. (2017a). *Switzerland - GDP Annual Growth Rate*. Trading Economics. April 10, 2017, from <http://www.tradingeconomics.com/switzerland/gdp-growth-annual>
- Trading Economics. (2017b). *Lebanon - GDP Annual Growth Rate*. Trading Economics. Retrieved April 10, 2017, from <http://www.tradingeconomics.com/lebanon/gdp-growth-annual>
- 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
- Tripathy, B., & Mishra, J. (2017). *A Generalized Framework for E-Contract*. *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*. IGI Global., doi:10.4018/IJSSMET.2017100101

## Compilation of References

- Tsankov, P., Andreeva, A., Yolova, G., Dimitrova, D. (2006). *Osnovi na publichnoto pravo*. Varna: Nauka i iкономика, IU-Varna.
- Türkmen, E., & Soyer, A. (2020). The Effects of Digital Transformation on Organizations. In *Handbook of Research on Strategic Fit and Design in Business Ecosystems* (pp. 259-288). IGI Global. doi:10.4018/978-1-7998-1125-1.ch011
- Uhl, L., & Gollenia, L. A. (2012). *A Handbook of Business Transformation Management Methodology*, Gower. SAP.
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (n.d.). Towards a global code of ethics for artificial intelligence research. In *Artificial Intelligence: the promises and the threats*. Retrieved from: <https://en.unesco.org/courier/2018-3/towards-global-code-ethics-artificial-intelligence-research>
- van Bommel, E., Edelman, D., & Ungerman, K. (2014). *Digitizing the consumer decision journey*. McKinsey Company. Available at: <http://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/digitizing-the-consumer-decision-journey>
- Vardan, M., Leyla, G., & Rinat, G. (2019). Design of Nano-scale Electrodes and Development of Avatar-Based Control System for Energy-Efficient Power Engineering: Application of an Internet of Things and People (IOTAP) Research Center. *International Journal of Applied Nanotechnology Research*. Advance online publication. doi:10.4018/IJANR.201901010
- Varkallo, V. (1978). *Ob otvetstvenosti po grazhdanskomu pravu*. Progres.
- Vasilev, A. (1993). *Zapravnata sashtnost na imushtestvenata otgovornost mezhdu stranite*. Zakon.
- Vasilev, A. (1997). Imushtestvvena otgovornost na rabotodatelya za vredy ot trudova zlopoluka i profesionalno zabolyavane. *Interyus*, 215.
- Vasilev, A. (2012). *Obezsheteniya po trudovoto pravootnoshenie*. Sibi.
- Vassileva, B. (2017). Marketing 4.0: how technologies transform marketing organisation. *Obuda University e-Bulletin*, 7(1), 47-56.
- Vassileva, B. (2016). Increasing cyber security competences through mission-based learning. *Proceedings of the International Conference on Human Systems Integration Approach to Cyber Security*, 189-204.
- Vassileva, B., & Zwilling, M. (2018). Hybrid Warfare Simulation-Based Learning: Challenges and Opportunities. *Information & Security*, 39(1), 220–234.
- Visser, J., Field, D., & Sheerin, A. (2015). *The Agile Marketing Organization*. The Boston Consulting Group. Available at: <https://www.bcg.com/publications/2015/marketing-brand-strategy-agile-marketing-organization>
- Vitorino, E. V., & Piantola, D. (2011). Dimensões da competência informacional (2) [Dimensions of information literacy]. *Ciência da Informação, Brasília*, 40(1), 99–110.

- Wallace, T. (2011). *Sales and operations planning: beyond the basics*. T.F. Wallace & Company.
- Wendell, W. (2010). *Moral Machines*. Oxford University Press.
- Wiener, N. (1948). *Cybernetics: Or control and Communication in the Animal and the Machine*. MIT Press.
- Wiener, N. (1954). *The Human use of Human Beings. Cybernetics And Society*. Doubleday Anchor Books.
- World Economic Forum. (2017). *The Global Risks Report 2017 12<sup>th</sup> Edition*. Available at: [http://www3.weforum.org/docs/GRR17\\_Report\\_web.pdf](http://www3.weforum.org/docs/GRR17_Report_web.pdf)
- Xavier, U. H. R., & Pati, B. P. (2012, November). Study of internet security threats among home users. In *2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)* (pp. 217-221). IEEE. 10.1109/CASoN.2012.6412405
- Yafushi, C. A. P. (2015). *A Competência em informação para a construção de conhecimento no processo decisório: estudo de caso na Duratex de Agudos (SP)* [Information literacy to construct knowledge in the decision-making process: a case study at Duratex Agudos (SP)] (Master's Dissertation). Retrieved from: <https://repositorio.unesp.br/handle/11449/126599>
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353. doi:10.1016/S0019-9958(65)90241-X
- Zaiane, S., & Ben Moussa, F. (2018). *Cognitive Biases, Risk Perception, and Individual's Decision to Start a New Venture*. *International Journal of Service Science, Management, Engineering, and Technology*. doi:10.4018/IJSSMET.2018070102
- Zandia, F., & Tavana, M. (2011). *A fuzzy group multi-criteria enterprise architecture framework selection model*. *Management Information Systems, Lindback Distinguished Chair of Information Systems, La Salle University*. Elsevier.
- Zhu, J., Huang, H., & Zhang, D. (2019). Big Tigers, Big Data: Learning Social Reactions to China's Anticorruption Campaign through Online Feedback. *Public Administration Review*, 79(4), 500–513. doi:10.1111/puar.12866
- Zwilling, M., Levy, S., Gvili, Y., & Dostal, P. (2020). Machine learning as an effective paradigm for persuasive message design. *Quality & Quantity*, 54(3), 1–23. doi:10.1007/11135-020-00972-0

## About the Contributors

**Bistra Vassileva**, PhD, works as Professor of Marketing at the University of Economics–Varna, Bulgaria. She graduated as Master of Science of Commodities. Since 1992 she is lecturing and consulting in the field of Marketing Research, International Marketing, Marketing Communications, TQM, Marketing Management. Dr. Vassileva was a visiting professor in Portugal, France, Germany, Spain, UK and a guest lecturer in Belgium. She implemented more than 15 international and national projects of different donors. During the last few years she took part in EU funded projects for various research issues and problems as an expert. She was a Marie Currie fellowship holder as a Senior researcher in 2007-2008 in Lodz, Poland. As a Director of the Centre for Innovation and Development at the University of Economics-Varna she is responsible for the organisation of research and ICT projects with different scope. Member of CIM, ESOMAR and EMAC. Her scientific interests are focused on nonlinear dynamics and network theory in the field of marketing, marketing analytics and digital marketing.

\* \* \*

**Andriyana Andreeva**, PhD, Master degree in Law, Professor in “Legal Studies” Department University of Economics – Varna since 1996, PhD Criminal Law since 2005. Associate Professor in Labour Law and Social security since 2011. Author of over 80 publications in the area of Labour Law and Social security.

**Rinat Galiautdinov** is a Principal Software Developer and Architect having the expertise in Information Technology and Computer Science. Mr. Galiautdinov is also an expert in Banking/Financial industry as well as in Neurobiological sphere. Mr. Rinat Galiautdinov works on the number of highly important researches as an independent researcher.

**Gait Klein** is a senior lecturer in the department of Economics and Business Administration at Ariel University. She received her PhD in Organizational Sociology

from The Hebrew University. Her research interests lie in the area of investments, decision making process and social psychology process during establishment of new ventures.

**Dietmar Köring** is an architect, researcher, and educator living in Cologne. He is head of the architectural research office Arphenotype, where he focuses on blurring the boundaries of different artistic disciplines. Dietmar was a research fellow at TU Berlin / CHORA City & Energy from 2012 to 2017 and has taught Digital Design at TU Braunschweig from 2010 to 2012, he was Guest Professor for Virtual Realities & Experimental Architecture at the University Innsbruck /Studio3 in 2011, Technology and Design Lecturer at the Cologne Institute for Architectural Design / C-I-A-D and visiting lecturer for digital design at the DeMontfort University Leicester. From 2011 to 2012 he was assistant professor for Smart City Concepts at the Technical University Cologne. He studied architecture at the University of Applied Sciences Cologne, the University of Western Sydney and at the Muthesius Academy of Fine Arts, where he graduated as in 2005 as Dipl.-Ing. (FH). Dietmar received his MArch in 2007 at the Bartlett School of Architecture University College London and his Dr.-Ing. at the Technical University of Berlin in 2018. Through his career he has worked internationally for offices such as Coop Himmelblau, Graft, 3deluxe and Andrew Wright Associates. His research has been awarded by the Jaap Bakema Fellowship / NAI and his works have been internationally published and exhibited. Dietmar has given international lectures, guest critiques and workshops.

**Arthur Kordon** is an internationally recognized expert in applying Artificial Intelligence to industry. His clients include large corporations and consulting companies in the United States, Japan, South Korea, Germany and Belgium. His current projects are solving business problems in predictive maintenance, smart energy cost reduction analysis, commodity price forecasting, and office space optimization. He also advises his clients on how to integrate AI into their organizations and trains their specialists. In his previous position as Scientific and Analytics Leader of the Artificial Intelligence Group at Dow Chemical – America's largest chemical company, Dr. Kordon introduced a number of new AI-based solutions that have improved and optimized manufacturing processes with enormous economic effect. He has a US and worldwide patent and has published three books, 15 book chapters and over 70 publications in the most prestigious journals and conferences in the field of applied AI systems.

**Dušan Lesjak** has been active as researcher, teacher and manager for more than 30 years. He is a professor of Management information systems (2000) and Management of Education (2010). His areas of expertise are e-business, e-learning

### ***About the Contributors***

and management and financing of higher education. He was a State Secretary in the Ministry of Higher Education, Science and Technology of Slovenia (2006-2008) and in recent years, Advisor to the Minister of High Education, Science, Culture and Sport, covering science and higher education. Before that he was a president of the Program Council for the Computerization of Schools, a member of the Commission of the Council of the RS for Higher Education, a member of the Council for General Education, a member of the Zois committee for awards and recognitions, etc. Already in 2006, 2 years before the official establishment of EMUNI, he chaired of the Project Group for EMUNI of the Government of Slovenia and has remained active by closely following the activities of the University ever since.

**Ilídio Manhique** is a Ph.D. Candidate in Information Science at Sao Paulo State University (UNESP) Brazil and Professor at Escola Superior de Jornalismo in Moçambique.

**Elaine Mosconi** is an Associate Professor At Université de Sherbrooke (UdeS), Canada.

**Selma Ottonicar** is a Ph.D. Candidate at Sao Paulo State University UNESP (Marilia campus) in the Information Science Postgraduate Program (PPGCI). On-line Tutor at Brazilian Association of Information Science (ABECIN). She holds a Masters in Information Science from UNESP where she received a CAPES scholarship. She is a member of the Information, Knowledge and Organizational Intelligence research group at UNESP and a member of the IntelliLab research group at Université de Sherbrooke (UdeS).

**Galina Yolova**, PhD, Master degree in Law, Professor in “Legal Studies” Department University of Economics – Varna since 1996, PhD Criminal Law since 2005. Associate Professor in Labour Law and Social security since 2011. Author of over 70 publications in the area of Labour Law and Social security.

# Index

## A

AI capabilities 1-2, 6, 14, 45, 56  
 AI ethics framework 54  
 AI policy 54, 57, 64  
 AI research clusters 1  
 AI strategy 54  
 Applied Competences 148-149, 155-156, 159  
 artificial intelligence 1-3, 6, 10, 13-16, 19-20, 22-25, 32-33, 35-37, 41, 46, 50-54, 58, 60, 64-65, 76-78, 81, 123, 179-182, 184, 188-189, 191-193, 195-202, 214-221, 231

## B

Bio-neuron 19  
 business applications 6, 9, 35-36, 42-43, 50-51, 60

## C

challenges 20, 32, 36, 48, 58, 63, 69, 100, 119-120, 122, 129, 149-150, 158, 160, 166, 178, 182, 193, 202-203, 205, 209-210, 212, 214, 221, 229  
 Covid-19 128, 144, 165-171, 173-174, 176-177  
 Critical Literacy 179  
 critical thinking 160, 179-180, 185, 191, 197-199, 201  
 cyber awareness 128-129, 132-133, 138  
 cyber readiness 171, 176  
 cyber security 73, 128-135, 137-153, 156-

163, 167-168, 170, 177, 208, 211, 225  
 cyber-crime 128, 130, 132, 166  
 cybersecurity 24, 66, 131-132, 138, 143, 150-151, 161-163, 165, 176-177, 180, 202, 208  
 Cybersyn 68, 71-79

## D

damage 11, 29, 81, 111, 114, 131, 150-151, 216-217, 219-221, 227, 231  
 deep learning 16, 41-42, 66, 183  
 digital avatar 20-21, 32-33  
 Digital Literacy 179  
 Disciplinary liability 216, 231

## E

ethical digital marketing 202  
 ethical guidelines for artificial intelligence 195  
 ethics 12, 19-20, 26-27, 32, 57-58, 69, 72-73, 76, 78, 180-181, 183-184, 188, 192, 195, 197, 200-201, 211  
 EU Legal Acts 231

## F

Fuzzy Logic Control 165

## G

geopolitical analysis 80-81, 85

## ***Index***

### **H**

Higher Education 124, 161, 180, 184, 198, 211

### **I**

ICTAIs 17

Industry 4.0 68, 70-72, 75-76, 78-79, 179-180, 198-199, 225

Information Evaluation 201

information literacy 179-184, 187-189, 191-201

Information Science 179-181, 188-189, 199, 201

Interdisciplinarity 201

Israel 128, 133-134, 142-143, 147, 165

### **L**

labor law 214-217, 222-225, 227

liability 150, 214-228, 231-232

Liability for Damages 219, 231

liability in labour law 227-228, 232

LIME (Local Interpretable Model-Agnostic Explanations) 17-18

### **M**

machine learning 1-2, 6-7, 11, 16-17, 19, 21, 26-33, 35, 40, 42, 46-48, 51, 135, 146, 178

Manager 82-83, 85, 89, 106-108, 127, 176

material liability 214, 226-227, 232

### **N**

Neural circuit 19

neural network 16, 29, 42

### **P**

Project 10, 26, 46, 49-50, 68, 70-71, 73-76, 78, 80-82, 85-87, 89, 91-95, 97-98, 101-108, 110-111, 115, 117-119, 123, 126-127, 199, 231

### **R**

REFRAC 148-149, 155-160

responsible AI 13-14, 36, 52, 54-55, 160, 202

risk management 85, 143, 220

### **S**

SHAP (SHapley Additive exPlanations) 17-18

simulation-based learning 148-149, 153-154, 158-160, 178

Slovenia 128, 133-134, 136, 147

social media 64, 166-167, 169, 202, 205-212

systematic literature review 179, 181, 188-189, 198

### **T**

types of legal liability 215-216, 223-224, 232

### **U**

universal basic income 68, 71

### **W**

WEAKNESSES OF AI 35, 41

weaknesses of human intelligence 35-37, 39