Premier Reference Source

Applied Ethics in a Digital World

Ingrid Vasiliu-Feltes and Jane Abigail Thomason

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Applied Ethics in a Digital World

Ingrid Vasiliu-Feltes
University of Miami, USA

Jane Thomason
University College London, UK

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

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

Library of Congress Cataloging-in-Publication Data

Names: Vasiliu-Feltes, Ingrid, 1971- editor. | Thomason, Jane, 1958- editor.
Title: Applied ethics in a digital world / Ingrid Vasiliu-Feltes and Jane Thomason, editors.
Description: Hershey PA : Information Science Reference, [2022] | Includes bibliographical references and index. | Summary: "This book highlights core dilemmas we are facing as a society in the digital era, reviews key domains that need to be addressed in designing and deploying proactive ethical and data governance programs, and offers potential areas for further research"-- Provided by publisher.
Identifiers: LCCN 2021023062 (print) | LCCN 2021023063 (ebook) | ISBN 9781799884675 (hardcover) | ISBN 9781799884682 (paperback) | ISBN 9781799884699 (ebook)
Subjects: LCSH: Information technology--Moral and ethical aspects. | Computer science--Moral and ethical aspects. | Artificial intelligence--Moral and ethical aspects. | Digital communications--Moral and ethical aspects. | Medical technology--Technological innovations--Moral and ethical aspects. | Applied ethics.
Classification: LCC T58.5 .A68 2022 (print) | LCC T58.5 (ebook) | DDC 174/.9004--dc23
LC record available at https://lccn.loc.gov/2021023062
LC ebook record available at https://lccn.loc.gov/2021023063

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All work contributed to this book is new, previously-unpublished material.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.

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

Mehdi Khosrow-Pour, D.B.A.
Information Resources Management Association, USA

ISSN:2328-1316
EISSN:2328-1324

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

- Cultural Influence of ICTs
- Human Development and Technology
- Human-Computer Interaction
- Technology and Freedom of Speech
- Technoself
- Human Rights and Digitization
- Technology and Social Change
- Information ethics
- Philosophy of technology
- Digital Identity

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

The Advances in 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. 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.

Titles in this Series

For a list of additional titles in this series, please visit: www.igi-global.com/book-series/advances-human-social-aspects-technology/37145

Technological Influences on Creativity and User Experience
Joshua Fairchild (Creighton University, USA)
Information Science Reference • © 2022 • 305pp • H/C (ISBN: 9781799843542) • US $195.00

Human Factors Issues and the Impact of Technology on Society
Heather Lum (Embry-Riddle Aeronautical University, USA)
Information Science Reference • © 2021 • 333pp • H/C (ISBN: 9781799864530) • US $195.00

Ubiquitous Technologies for Human Development and Knowledge Management
Hakikur Rahman (Institute of Computer Management and Science, Bangladesh)
Information Science Reference • © 2021 • 396pp • H/C (ISBN: 9781799878445) • US $195.00

Technological Breakthroughs and Future Business Opportunities in Education, Health, and Outer Space
Angus Hooke (Australian Institute of Higher Education, Australia)
Business Science Reference • © 2021 • 353pp • H/C (ISBN: 9781799867722) • US $225.00

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

Examining the Socio-Technical Impact of Smart Cities
Fenio Annansingh (York College, City University of New York, USA)
Information Science Reference • © 2021 • 231pp • H/C (ISBN: 9781799853268) • US $190.00

Human-Computer Interaction and Technology Integration in Modern Society
Hakikur Rahman (Ansted University Sustainability Research Institute, Malaysia)
Engineering Science Reference • © 2021 • 347pp • H/C (ISBN: 9781799858492) • US $195.00

Information Technology Applications for Crisis Response and Management
Jon W. Beard (Iowa State University, USA)
Information Science Reference • © 2021 • 294pp • H/C (ISBN: 9781799872108) • US $195.00

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

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

Table of Contents

Detailed Table of Contents

Section 3
Digital Ethics in Health Care

This chapter discusses digital equity through the lens of the digital divide. While the digital divide is as old as information communication technology itself (ICT), the COVID-19 health crisis renewed a strident interest in exposing the significant gap that still exists after close to 30 years. The digital divide then is first contextualized within the coronavirus pandemic to illustrate how inequities came further to the forefront of people's agenda. It then moves to discuss the digital divide defining the complex term and offering critical data to illustrate the areas of the world most impacted by this unfortunate reality. Different organizations and groups have made significant moves to narrow the digital gap. These strategies are discussed next. None of these groups will be fully successful if, as will be argued, they are not concerned with digital equity. Finally, the chapter makes some critical observations on future challenges facing ICT vis-à-vis the digital divide.

The advances in biotechnology and computer and data sciences opened the way for innovative approaches to human healthcare. Meanwhile, they created many ethical and regulatory dilemmas such as pervasive global inequalities and security and risk to data privacy. The assessment of health technology is a systematic multidisciplinary process that aims to examine the benefits and risks associated with its use including medical, social, economic, and ethical impacts. It is used to inform policy and optimize decision-making. The advance of technology is creating significant challenges to healthcare regulators who strive to balance patient safety to fostering innovation. The FDA and EMA are modernizing their regulatory approaches to foster innovation in digital technology and improve safety and applicability to patients. On the other hand, data analytic technologies have been introduced into regulatory decision processes.

As organizations steadily adopt remote and virtual capabilities, informed consent processes are increasingly managed by digital technologies. These digital methods are generating novel opportunities to collect individuals' permissions for use of private information but are blurring traditional boundaries of consent communication and documentation. Therefore, the rapid growth of digital technologies used for informed consent as well as the sheer volume of data resulting from electronic data capture are generating complex questions about individual engagement and data practices. This chapter presents emerging risks, benefits, and ethical principles about digital informed consent methods and technologies. For the areas where digital informed consent creates ethical uncertainties, ethical guidelines and user-design recommendations are provided.

Section 4
Digital Ethics and New Realities

Telemental health (TMH) is considered by many to be the future of mental healthcare, with some claiming that these methods should replace more traditional approaches. Early teletherapeutic initiatives demonstrate an immediate set of benefits for patients including improved access to care, reduced costs, better schedule flexibility, greater environmental familiarity, and higher rates of patient engagement. Notable limitations to TMH include enhanced privacy concerns, the variable digital literacy of certain populations/persons, and technological instability. However, other limitations regarding therapeutic relationships, experiences, and settings have gone undertheorized and are not sufficiently represented in the current research. This chapter surveys these considerations and argues that digital medical interventions are unable to effectively replicate the same degree of 'contact' and 'intimacy' available in physical care; providers should therefore be cautious in wholly replacing in-person methods or in implementing a standalone paradigm of digital care.

Social media is a mega-industry built by systematically monetizing the exploitation of human emotions, reactions, and biases. The authors explain how this industry became so profitable by creating a fear of missing out (FOMO) to command our attention, blending news and content in one feed to keep users 'in-app', and using powerful algorithms to promote more provocative posts, filter content, and trigger the reward centres of our brains. The authors examine how decentralized technologies, including cryptocurrencies, tokenization, and blockchain are being developed and deployed into new social media applications. The authors speculate on how these blockchain-backed startups could challenge the status quo and appeal to new expectations of user privacy, tighter regulation, and a more equitable monetization system.

This chapter focuses on aspects of digital ethics in technology and investments and covers the background of the growing surface area of the technology and its concerns related to data; it provides an overview of the data-related challenges and their ethical uses by organizations and people. It also covers emerging technology like artificial intelligence and its impact on ethical challenges. It provides an overview of the potential ethical challenges of technology that investors should consider and focuses on the criticality of the framework requirement and its implementation within businesses to make the right decisions. The author lays out the view on ethics and regulations and why companies should commit to ethical practices for their growth.

This chapter will highlight the importance of transforming our conceptualization of business ethics in the digital era and the opportunities related to an optimal design of sustainable digital business ethics programs in this new hyper-connected, hyper-automated digital world. The complex issues of this revised business ethics model will be addressed from three perspectives: corporate governance, leadership, and society. The sections related to corporate governance will highlight the operational challenges when aiming to incorporate ethics into the boardroom's DNA and will emphasize the sustainability imperative ethical business leaders are facing in this digital era. This chapter will also posit that by adopting a design thinking approach for business ethics in this digital era, we can leverage all the benefits offered by emerging technologies and scientific advances while maintaining a human-centric stance.

The rapidly growing field of Blockchain and Decentralised Finance (DeFi) has the potential to transform many aspects of the financial world. It offers an abundance of opportunities to reduce costs, increase transparency and reduce the need for middlemen in financial services. While this promise of automation and decentralization is attractive, it is important to consider the potential for inadvertent or deliberate automation of unethical conduct at scale. Ethical questions involve the consideration of conflicting moral choices and dilemmas. Blockchain creates ethical dilemmas for developers, investors, consumers, and regulators at the technology, application, and societal levels. This chapter provides a perspective on the emerging field of DeFi and Blockchain in financial services, a reflection on the ethical questions that arise, how they are being addressed, the key issues, and further research needed in this growing field of Blockchain ethics. The objective of the chapter is for students, developers, CEO's and governments to appreciate the moral and ethical issues being uncovered in the course of the development and deployment of Blockchain technology. Blockchain, as with all technology, is a tool and is as beneficial and useful as the care that is taken to make it. There remains a need to ensure that Blockchains are built and deployed with due concern for ethics.

The ethical risks which emerge from the cross section of artificial intelligence, extended reality, and geographic information systems could be examined in two broad categories of environmental and user-centric interactions of human beings with AI-curated mixed realities. These categories resonate with the capacity of AI to significantly impact the efficient application of extended reality technologies, while utilizing geodata and behavioral modelling to alter and transform experiences. While regulatory frameworks are catching up with the rights of users in the digital economy, the recently accelerated growth of immersive technologies provides further scenarios and use cases, which ought to be considered for their capacity to amplify biases, produce alternative realities, and affect human emotions.

Foreword

The scientific and technological innovations that are driving the development of expert systems and, perhaps one day, genuine 'artificial intelligence' are rapidly bringing humanity to the brink of civilizational change. This is no idle claim. When combined with other technologies in the fields of robotics, synthetic biology and nano-manufacturing (to name but a few), it is possible to envisage a world in which the vast majority of humanity enjoy access to the material necessities of life without the burden and drudgery of toil. Of course, this 'utopian' vision has been advanced by others before. However, those visions have typically depended on a fictional account of what might be the fruits of human ingenuity. The difference this time, is that those fruits are already being harvested.

Of course, imagined futures also include far darker, 'dystopian' possibilities. There is no need to canvass, here, their general lines. It is, instead, it important to note that the fears that they inspire are already gaining traction. If they occupy a place in the minds of enough people, then this could lead to otherwise useful developments being restrained or abandoned.

So, where does the difference lie in terms of which possible future will emerge?

As I have argued elsewhere:

Technical mastery divorced from ethical restraint is at the root of all tyranny.

We have seen the truth of this maxim expressed in events throughout human history: the weaponization of nuclear physics, the application of superb engineering to create maximally efficient death camps, the list is long.

Too often, the malicious use of science and technology has been enabled by the absence of even the most minimal source of ethical restraint. That is, people have not even realized that there were fundamental ethical issues to be considered. Only after the proverbial 'horse had bolted' – and the damage had been done – would someone pause to ask, "how was this possible? What were we thinking? Oh, the horror!".

It is against this background – of both risk and opportunity – that we should welcome this edition of essays on the general topic of ethics and artificial intelligence (AI). A book of this kind is an invitation not only to reflection but to responsibility. It allows us to pause, if only for a moment, to consider not only if something *can* be done but also whether it *should* be done.

Such a pause is, in itself, potentially revolutionary – especially if the opportunity to reflect is built into the structure of scientific and technological developments in this area. For my part, I think it essential that the topics covered in this book be used to inform the structure of development and application. Indeed, I think that emerging technologies (such as blockchain) should be used to track the ethical 'provenance' of AI systems – from the moment of inception, to design, application and modification.

Each stage should require a statement of the core purpose, values and principles that have been taken into account and most importantly, how they have been given practical effect in the design, development and implementation phases. Any subsequent change should also register any variation from the patter of ethical design. Finally, all of this should be open to assurance by disinterested professionals (tech auditors) who van verify that the stated ethics framework is being expressed in functional terms.

Why Argue for such a regime? It's not simply a response to the growing power of the technology (although that is a factor). Rather, it's more to do with the unusually 'opaque' character of systems like unsupervised neural networks. We can know the initial conditions, we can know if the system works but, in principle, we cannot precisely describe every step in the process (or overlapping processes) that produce a successful outcome. Not knowing the detail necessarily wakens the 'chain of *accountability'*. This the need for a higher degree of *responsibility* on the part of those who create and deploy these technologies – from the outset.

The authors who have contributed to this book have started at 'ground level' launching the trajectory of their shared perspectives from a solid grounding in human rights. From there, the 'sky is the limit' – reaching areas where authonomous expert systems may be directing the affairs of human beings – in effect acting as our 'boss'.

As others have noted, in the past, we should recognise that human beings are no less susceptible to the effects of bias, error, fallible judgement than the machines we are developing. Already, there are some functions where AI significantly outperforms when compared to humans attempting the same tasks. For example, AI exceeds the capacity of all humans in diagnosing, quickly and accurately, some forms of cancer. However, I would argue that only another human can place their hand on another's shoulder to convey the results of a diagnosis – especially if the most likely consequence is otherwise premature death. The difference lies in the fact that only another human being knows what it is to be mortal and therefore can convey, with genuine sympathy, the prospect of death. Yes, an adequately trained system can offer a perfectly convincing performance of such a conversation – but it will always lack knowledge of a truth that makes the world of difference.

There is some debate about whether or not AI will be able to handle genuine ethical dilemmas when, in fact, values and principles (and all that flow from them) are so perfectly balanced as to make the 'right' decision impossible to discern. Humans encounter this all the time – and often manage to resolve such dilemmas by resorting to an option that is a pure act of creation; something so unexpected as to be unpredictable. I mention this final question because it brings us back to the central purpose of this book.

It is a point so obvious as to risk being ignored … this book has been written by humans for humans to read. That is, it begins with a recognition of the fact that we are entirely, inexcusably responsible for the things we make and use. So, let's hope we absorb the impact of that lesson. Let's hope that we think before we act; that we anticipate (as best we can) all that might follow from the choices we make.

As noted at the outset, I believe that AI will transform our world. With books like this at hand, I hope for a better future arising out of a transformation that is both just and orderly. Get the ethics right, and we will be able to look back at this time of decision without regret.

Simon Longstaff
School of Business, University of New South Wales, Australia

Preface

OVERVIEW

As advances in disruptive technologies transform politics and increase the velocity of information and policy flows worldwide, the public is being confronted with changes that move faster than they can comprehend. There is an urgent need to analyze and communicate the ethical issues of these advancements, as well as offer potential pathways to address the moral dilemmas we are facing as a digital society. In a perpetually updating digital world, data is becoming the dominant basis for reality. This new world demands a new approach because traditional methods are not fit for a non-physical space like the internet.

Applied Ethics in a Digital World provides an analysis of some of the critical ethical questions raised by modern science, technological advancements, and the fourth industrial revolution and explores how to harness the speed, accuracy, and power of emerging technologies in policy research and public engagement to help leaders, policymakers, regulators, academicians, and the public understand the impact that these technologies will have on economies, legal and political systems, and the way of life.

WHERE THE TOPIC FITS

The many academic areas covered in this publication include, but are not limited to:

- AI Ethics
- Artificial Intelligence (AI)
- Blockchain
- Business Ethics
- Cyber Security
- Decentralized Technologies
- Digital Divide
- Digital Equity
- Digital Ethics
- Digital Mental Health
- Digital Rights Management
- Digital-Informed Consent
- Emerging Health Technologies
- Ethical Risks

- Extended Reality (XR)
- Geographic Information Systems (GIS)
- Human Rights
- Human-Machine Interaction
- Rights of the Child
- Self-Aware Artificial Intelligence
- Social Media Megacorporations
- Technology and Investments
- Translational Ethics

TARGET AUDIENCE

Covering a broad range of topics such as artificial intelligence (AI) ethics, digital equity, and translational ethics, this book is a dynamic resource for policymakers, civil society, CEOs, ethicists, technologists, security advisors, sociologists, cyber behavior specialists, criminologists, data scientists, global governments, students, researchers, professors, academicians, and professionals.

AN INTRODUCTION TO THE BOOK

The book's chapters are organized in five sections. The first section highlights the need to acknowledge and accept translational ethics as a new discipline and the powerful impact digital ethics has on socio-political decision-making. The second section calls attention to the complex and far reaching impact digital ethics can have on our efforts to uphold human rights and maintain equity in a digital society. The third section is focused on nuances related to various domains in digital healthcare ethics and the fourth section aims to showcase how digital ethics must harmonize with business and investments ecosystems, as well as shed light on future realities we must proactively address to avoid negative ethical consequences. In the closing section, we point to future research required to dive deeper into some of the ethical divides described and tackle the yet untouched issues on the digital ethics agenda.

In Chapter 1, "Introduction and Importance of Digital Ethics as an Emergent Discipline," David Danks argues first that ethical decisions and considerations are ubiquitous within the creation of digital technology. Ethical analyses cannot be treated as a secondary or optional aspect of technology creation. He then argues that this research must take the form of translational ethics: a robust, multi-disciplinary effort to translate the abstract results of ethical research into practical guidance for technology creators. He calls for a meaningful digital ethics conversation that actually leads to technology that better supports our values involves the conversion of these principles into useful, tangible practices, noting that just as translational medicine converts biomedical research into clinically useful guidance, we need a translational digital ethics to yield useful changes to our current technology creation pipeline.

In Chapter 2, "The Role of Tools in Advancing Ethical AI: Opportunities and Limitations," Ivana Bartoletti argues that a dimension of fairness in algorithmic decision-making is necessary, however, to achieve equity and transparency, this needs to take the form of an active choice mobilized by public awareness, accountability measures, and a reformed public policy to address concerns arising from flawed automation and public relation (PR) slogans. She makes the point that fairness may not often be the optimal financial solution for an organization, as it requires the manipulation of the data and the AI artifact that processes it. For this reason, choosing to adopt fairness in algorithmic decision-making is an inherently political, social and ethical choice. Accordingly, she argues that AI artifacts should be conceptualized as socio-technical systems, encompassing both a technical and social approach. Bartoletti proposes that the deployment of flawed automation will inevitably harm trust in technology, perpetuating existing inequalities at the risk of hindering the fundamental rights of privacy, human dignity, freedom, and autonomy. She evaluates the ways in which such tools, when adopted in isolation and without a standardized approach to fairness, can lead to further damage by removing the wider socio-political dimension of algorithmic decision-making.

In Chapter 3, “Ethics, Digital Rights Management, and Cyber Security,” Ali Hussain explores Digital Rights Management (DRM) practices for creating better and ethically safe online spaces. He discusses state of the art DRM paradigms, critically discusses their technical performance, flexibility and immutability challenges, and critically analyzing the cloud technology stack and ethical features. The integrity of online content lacks some features compared to the real adoption and deployment challenges, such as interoperability, internationalization, and internet policy and argues that research on this issue is needed. He highlights the need to find new and robust and tamper proof authentication and authorization methods for safer and secure internet. He concludes that the development of an advanced version of the DRM System suitable for cloud computing would enable researchers to protect online content more effectively, with greater content security.

In Chapter 4, “Keeping the U.N. Convention on the Rights of the Child Relevant in the Digital Age,” Susan Zinner notes that the internet offers opportunities for children around the world to connect with each other and to learn from each other. Until very recently few people felt that children deserved any rights at all and some adults have been unwilling to acknowledge the cognitive maturity of others. Since children mature emotionally and cognitively at different rates—just as their physical developmental rates vary—many societies have traditionally opted to err on the side of caution and limit their exposure to individuals, substances and media deemed risky. The central role that the digital world now plays in the lives of all global citizens today, including children could not have been imagined. Zinner explores how the UNCRC would likely address the right to digital access by children, how adults in the lives of children should balance the protection and participation rights of children seeking to exercise these rights, what guidance the “evolving capacities” standard in the UNCRC provides in the context of minors exercising participation rights, how to respond to concerns about unequal digital media access, issues of good child citizenship involving digital media and international laws and reports that may provide some guidance in resolving these issues.

Zinner explores participation and empowerment and how to ensure that children have the requisite tools to move beyond participation so that they are empowered to make decisions which have meaningful consequences. Zinner concludes digital access is now necessary for basic communication, work, entertainment and other purposes. In order to meet the needs of children around the world today, parents, legislators and those who simply care about the future of children, need to ensure that children have the tools to allow their voices to be heard.

In Chapter 5, “Instructing AI Ethics and Human Rights,” Miller and Muhammet Demirbilek considers emerging issues of AI and the current literature on AI ethics and human rights teaching. He explores teaching methodologies to explain AI ethics and human rights in K-12 learning environments. Particular emphasis is made on a survey of existing ethics teaching methodologies and how to adopt existing teaching strategies into AI ethics teaching in order to improve student understanding on AI ethics and human rights. He emphasises the educational aspects and the need to teach ethical and moral issues caused by AI powered technologies and to prepare future generations with AI ethics competencies to critically reflect on their learnings in their future jobs.

In Chapter 6, "Digital Equity," Patrick Flanagan discusses digital equity through the lens of the digital divide, using the coronavirus pandemic to illustrate how digital inequities came to the forefront of people's agenda during this period. It then moves to discuss the digital divide defining the complex term and offering critical data to illustrate the areas of the world most impacted. To resolve the threat of the continuance of the virus, vulnerable communities were targeted to ensure that these people already on the fringes did not suffer anymore than necessary. Government and health officials made strident efforts in the name of equity to resolve the occurrence of the virus in these edges of society. He proposes that to esteem the value of equity it is necessary to address the digital needs of those on the peripheries of the global village with greater vigilance.

In Chapter 7, "Ethical and Regulatory Challenges in Emerging Health Technologies," Samia Rizk considers the advances in biotechnology sciences such as genomics, neuroscience, synthetic biology, and nanoscience combined with the rapid developments in computer and data analytics, and the challenges of the increasing complexity of ethical issues associated with modern health technology. Health technology is rapidly evolving and transforming the healthcare system. Data-driven systems are particularly changing traditional healthcare delivery. The gap in knowledge and expertise between these innovative interventions and the existing healthcare contexts for both givers and users, creates several ethical, regulatory, and economic challenges which could destabilize trust in the safety, fairness, and effectiveness of the healthcare system. She argues for the need for specialized expertise and wider scope of analysis and a new subfield of ethics termed technology ethics or "techno-ethics" which deals with the framing of principles and methods to guide technology implementation and use.

In Chapter 8, "Ethical Benefits and Drawbacks of Digital Informed Consent," Wendy Charles et.al. considers how the emergence of digital methods has created opportunities to both enhance ethical protections and detract from intended protections. The chapter primarily focuses on informed consent processes that require a complex exchange of information, such as informed consent for healthcare delivery and participation in human subjects research. The authors' stance throughout this chapter is that the adoption of emerging digital technologies for informed consent does not alter ethical principles, since these retain the moral compass shaped by societal values rooted in autonomy and justice. Digital informed consent methods instead require adaptations of consent processes and appropriate uses of permissioned data to adhere to ethical principles. She argues for an individual-centered process which places the well-being and interests of the individual at the forefront and uses the technology to overcome a person's limitations to make a truly informed decision. Organizations designing digital informed consent methods must assume responsibility for obtaining agreement in an ethical manner. They can begin by creating ethical frameworks that delineate an organization's values, detail ethical assessments, and risk mitigation strategies that create an environment more likely to promote a trustworthy agreement and integrity with data protections and use. As digital technologies evolve, laws, regulations, and guidelines must change accordingly to maintain the individual at the center.

In Chapter 9, "Going Telemental: Contact and Intimacy in Digital Mental Health," Shaun Respess looks at telemental health (TMH) which offers unique techniques for persons to interact without needing to be physically present. He examines the long-term advantages and fit alongside of conventional therapeutic methods. The author applies elements from care ethics to explore and critique the sustainability of TMH, arguing that such services can compromise the quality of care even while providing several benefits. Furthermore, the author suggests that critical examinations of 'contact', 'intimacy', and embodied spaces should be crucial features of therapeutic assessment, including evaluations of TMH. He concludes that whether digital technology will be an enabling or restrictive force moving forward is at present unclear, but it is certainly an intriguing source of controversy and exploration.

In Chapter 10, "Social Media: How a Mega-Industry Was Built by Systematically Monetizing the Exploitation of Innate Human Emotions and What, if Anything, Can Change This," Foster Fletcher tackles social media and its systematic exploitation of human emotions, reactions and biases, blending news and content in one feed to keep users 'in-app' and using powerful algorithms to promote more provocative posts, filter content and trigger the reward centres of our brains. He then explores the potential of decentralised technologies, into new social media applications, providing new expectations of user privacy, tighter regulation and a more equitable monetisation system. He speculates that social networks that successfully protect their users' privacy, reduce hate speech, block bots, reward contributions, and respect personal data have a future. He concludes that people will reject the current social media paradigm and opt for a decentralised, distributed and equitable future.

In Chapter 11, "Digital Ethics in Technology and Investments," Ritesh Jain focuses on aspects of Digital Ethics in technology and the growing risks related to data. He provides an overview of the data-related challenges and its ethical uses by organisation and people, as well as emerging technology like AI and its impact on ethical challenges. Based on this he examines the potential ethical challenges of technology that investors should consider and focus on the criticality of the framework requirement and its implementation within businesses to make the right decisions. Finally Jain lays out the view on ethics and regulations and why companies should commit to ethical practices for their growth. Organisations need to continuously educate ethical decision-making, reinforce, monitor, and empower employees to question any potential unethical issues.

In Chapter 12, "Business Ethics," Ingrid Vasiliu-Feltes highlights the importance of transforming our conceptualization of business ethics in the digital era and the opportunities related to an optimal design of sustainable digital business ethics programs in this new hyper-connected, hyper-automated digital world. The complex issues of this revised business ethics model are addressed from three perspectives: corporate governance, leadership and society. The sections related to corporate governance highlight the operational challenges when aiming to incorporate ethics in the board room's DNA and will emphasize the sustainability imperative ethical business leaders are facing in this digital era. Throughout the chapter an emphasis is placed on the crucial importance of ethical leadership in this digital era and the unique characteristics required for long term success.

In Chapter 13, "Ethical Challenges for Blockchain and Decentralised Finance (DeFi)," Thomason discussed the rapidly growing field of Blockchain and Decentralised Finance. She considers the potential for inadvertent or deliberate automation of unethical conduct at scale and highlights ethical dilemmas for developers, investors, consumers and regulators at the technology, application, and societal levels. The chapter provides a perspective on the emerging field of DeFi and Blockchain in financial services, a reflection on the ethical questions that arise, how they are being addressed, the key issues, and further research needed in this growing field of Blockchain ethics.

In Chapter 14, "Ethical Risks in the Cross-Section of Extended Reality (XR), Geographic Information Systems (GIS), and Artificial Intelligence (AI)," Manolova explores the expansion of extended reality, geographic information systems and artificial intelligence. She makes the point that the mainstreaming of data science has been accompanied by expanded daily hours in front of screens and emergence of a brand-new internet minute, which contains within itself not only an addicting human-digital symbiosis, but fundamental dependence and emotional investment into digital solution as a bridge between families, friends, communities, and societies. She notes that the generational gaps in digital product consumption are constantly growing in sync with emerging technologies as the ways in which society consumes information (streaming, social networks, digital realities, immersive environment) and participates in economic exchanges (freelancer, influencer economics) continue to evolve along with cutting edge technologies, which were not available even five years ago. The addictive nature of the internet word transforming the internet minute into the first generation of internet lifetimes, where significant milestones in the human experience could exist almost entirely online. Large demographics already buy in the notions of reality TV, steaming entertainment, social media as forms of escapism and the lines between the real world and the digital one is becoming blurrier with the convergence between social networks and geo locations. Manolova makes the point that data science and innovation practitioners within the field of AI, need to maintain integrity in the face of quick and easy business solutions and to treat every user of the emerging technologies as a multi-dimensional human being and not a number on a dashboard, removed from implications outside of the digital domain.

In Chapter 15, “My Boss Is an Algorithm: Discussions on How to Best Prepare Students on the Ethics of Human-Machine Interactions at Work,” Cynthia Montaudon Thomas explores how employers can address whether or not algorithms will be able to treat people fairly. She reports on a study centered on a private higher education institution in central Mexico. General conditions regarding stress and working hours in the country are described to create the general background of the study, along with two significant regulations that legislate psychosocial risks and remote work. The population considered were full-time faculty members who had moved their activities online. Results were analyzed as a whole and later on divided according to gender to determine whether there were significant differences in terms of burnout syndrome in faculty. She concludes that the digital literacy that is instrumental to educational, working, personal, social lives, and algorithmic literacy needs to include the ethical design of algorithms to make sure that decisions are fair and transparent . One of the critical issues is how to prepare students for a world of machine learning in which algorithms can be used to make work more efficient, and at the same time, ethical, trustworthy, and fair.

In Chapter 16, “Self-Aware Machines,” Robin Craig discusses the ethical questions in terms of both how to ensure that self-aware machines will act in the interests of human beings and what rights such machines could have themselves. The chapter describes the general principles of AI ethics; the likelihood of self-aware machines being created; the implications of imposing pro-human ethical constraints upon them; the critical need this creates for an objective ethical system; a proposed objective ethical system, including its implications for the ethical relationship between humans and artificial minds and the further implications this holds for future research in both ethics and computing.

In the conclusion, Thomason and Vasiliu-Feltes provide a summation of the collection of articles which explore the many areas that ethics need to be deployed proactively and mindfully in order to ensure that we build technology that has ethics embedded by design. They discuss how to establish global ethical governance, the application of design thinking in designing ethical deployments of emerging technologies and the complex relationship between political power and ethics, as well as future research issues for ethics and technology. As highlighted by various chapters throughout the book, emerging technologies are strong drivers of innovation and transformation, however, they can also act as gateways for unintended negative ethical breaches or as powerful weapons when used by those with malicious intent. They conclude that as a society, we must focus on the opportunity to seek alignment of digital ethics programs with compliance, regulatory and ESG-conscious policies. This will require a high degree of synchronization and harmonization among all global stakeholders to achieve long-term success.

CONCLUSION

The purpose of this book is to provide objective information regarding the emergent field of applied ethics in a rapidly accelerating digital world. Thus, the authors see that this book is but an introduction to digital ethics through the perspectives of global practitioners and scientists.

The adoption of emerging digital technologies does not alter ethical principles, since these retain the moral compass shaped by societal values rooted in autonomy and justice. Digital ethics are not different from conventional ethics, but it is the potential for inadvertent or deliberate automation of unethical conduct at scale that highlights ethical dilemmas for developers, investors, consumers and regulators at the technology, application, and societal levels. New and powerful technologies demand a new approach and one that is co-developed with industry.

Ethical decisions and considerations cannot be treated as a secondary or optional aspect of technology creation. There is a need to translate the abstract results of ethical research into practical guidance for technology creators, with digital ethics that is fully integrated into the practices of digital technology creation. Technologists need to be able to collaborate closely with people from other disciplines to translate ethical insights into actionable, practical changes in our people, processes, policies, and partnerships. Inherent in this will be political, social and ethical choices.

The book is pragmatic, as academics, technologists, industry and government are coming to understand the demands of "technoethics." The chapters represent a snapshot in time and the velocity of change is dramatic. In the time it takes you to read this book there will be changes to protocols, nomenclature, approaches and methodologies. We hope that what we have compiled is a foundation for others to build on.

Acknowledgment

We would like to express our gratitude to a wonderful group of professionals who generously shared their expertise and insights during our Exponential Digital Ethics Series and inspired us to embark on this editorial journey. Their reflections and opinions generated increased awareness among our audience, pointed out numerous ethical challenges we are facing in this new digital world and encouraged us to take a deeper dive to learn more about viable solutions that applied ethics can offer.

We are extremely grateful to all authors for their important contributions and to our publisher for the support provided during this process.

Kenneth Goodman, PhD, Director, Institute for Bioethics and Health Policy

Matthew Rosenquist, CISO at Eclipz.io Inc. - Formerly Intel Corp, Cybersecurity Strategist, Board Advisor

Reid Blackman, PhD, CEO Virtue | AI Ethics Advisor

Cara Lapointe, PhD, Co-Director of the Johns Hopkins Institute for Assured Autonomy

Ivana Bartoletti, Global Data Privacy Officer, Wipro | Visiting Policy Fellow, Oxford Internet Institute | Author

Section 1

Introduction to General Concepts

Chapter 1
Digital Ethics as Translational Ethics

David Danks
University of California, San Diego, USA

ABSTRACT

There are growing calls for more digital ethics, largely in response to the many problems that have occurred with digital technologies. However, there has been less clarity about exactly what this might mean. This chapter argues first that ethical decisions and considerations are ubiquitous within the creation of digital technology. Ethical analyses cannot be treated as a secondary or optional aspect of technology creation. This argument does not specify the content of digital ethics, though, and so further research is needed. This chapter then argues that this research must take the form of translational ethics: a robust, multi-disciplinary effort to translate the abstract results of ethical research into practical guidance for technology creators. Examples are provided of this kind of translation from principles to different types of practices.

INTRODUCTION

As our world becomes increasingly digital, we must ensure that we do not lose our ethical compass. Algorithms are now frequently used to determine the allocation of critical resources, in some cases even making literal life-and-death decisions. Our lives are measured, collected, analyzed, and stored, thereby reducing us in some cases to simply a set of numbers. We carry digital devices that track our every location, volunteer information to improve our online experiences, and extend our minds, families, and communities through digital means. But much, perhaps all, of these transformations have been designed, implemented, and dictated primarily by technological and economic demands. We risk the creation of digital technologies that serve the values and interests of the few, rather than the values and interests of the many. The recent calls for digital ethics are, at their heart, an effort to return our focus to humans and our values. How do we have the right digital technology for us, and how can we achieve it in practice?

One barrier to the development and implementation of digital ethics has been uncertainty or misunderstanding about the scope and nature of ethics itself. For some people, ‘ethics’ refers to their personal

DOI: 10.4018/978-1-7998-8467-5.ch001

or cultural beliefs and principles. For others, 'ethics' is a matter of law, regulation, and compliance. Or perhaps 'ethics' is understood as an entirely relativistic domain where there truly are no right and wrong answers, or even better and worse answers. Along a different dimension, some people view 'ethics' as a purely negative enterprise that provides only restrictions or constraints (e.g., "killing someone is not morally acceptable except in certain extreme circumstances"). Alternately, one could conceive of 'ethics' in terms of positive principles (e.g., "you ought to have freedom of expression"). Of course, the reality of ethics is more complex than either of these possibilities, even if only because many positive principles presuppose restraint or constraint on others. For example, if I have freedom of expression, then others are prevented from stopping my speech (unless there are compelling counter-reasons).

In light of these disagreements about the exact nature of ethics, one might wonder whether any progress is possible at all, let alone in the specific case of digital ethics. Perhaps surprisingly, though, it can sometimes be easier to gain ethical insights in specific domains, rather than always trying to operate at an extremely high level of generality. As we briefly discuss below, 'digital ethics' is not properly understood as the "mere" application of well-established ethical principles to the particular case of digital technologies. Rather, we need to develop ethical principles, guidelines, and practices that are specific to the digital ecosystem. Even with this narrowed focus, though, we might still wonder "what is ethics?" For the purposes of this chapter, I will adopt a relatively simple characterization, with the full awareness of its limitations when trying to do highly abstract and general ethics.

I will understand 'ethics' as primarily concerned with two questions: (1) what values ought we have? and (2) given our values, how ought we act? Neither of these questions will have a unique answer. Ethical reasoning and analysis almost never determine our values; some of my interests and needs are specific to the peculiarities of my situation and life. And even if we know the values and interests of all of the relevant people, there will typically be multiple actions that are morally acceptable. Moreover, I acknowledge that these specific questions lack the subtlety that we normally expect and require from ethics, but they are nonetheless valuable in focusing our attention in productive directions. In particular, they place the focus of ethics squarely where it should reside—on the people who are impacted by the digital systems, technologies, and actions.

One of the background themes of this chapter is that digital ethics is intimately connected with the practices of digital technology creation. We need to ensure that our digital ethics is grounded in the actual challenges that we face in digital spaces, rather than proceeding at an unhelpfully abstract level. The next section thus examines the locations of digital ethics: where do ethical (and societal) questions arise in the digital technology pipeline? The following section then considers how we can translate our insights about people's values and interests into novel and improved practices. In that regard, this chapter can be understood as setting the stage for many of the other chapters in this volume. The practices of digital ethics should not be "mere" implementations of abstract ethical principles. Rather, they should be effective, grounded, focused practices that ensure our values and interests are realized through digital technologies, even if 'ethics' is not a term that appears in many of those practices.

UBIQUITY OF DIGITAL ETHICS

One common (though incorrect) claim is that (digital) technologies are not themselves the subject of ethical analysis. More specifically, this claim—sometimes called the "neutrality thesis" (e.g., Dotan, 2020; Friedman & Hendry, 2019; Simon, 2017)—says that the proper subject of ethical analysis is the *use* of technology, not the technology itself. For example, it seems quite odd to ask about the ethical values of a hammer, though we can surely inquire about the ethics of *using* a hammer in various ways. More generally, we do not normally think or talk about the ethics of particular objects or artifacts, but only the ethics of different uses of those objects or artifacts (though see, e.g., Winner, 1980). One might reasonably think that the same should hold for digital technologies. However, the neutrality thesis cannot hold for digital technologies for at least two distinct reasons.

The first problem with the neutrality thesis is that, in contrast with relatively inert physical artifacts, digital technologies sometimes make ethical decisions themselves. Autonomous and semi-autonomous technologies are increasingly being used across society, and these systems typically have the capability to plan, decide, and act in the world. These systems make ethical choices throughout their operation; in some cases, these choices are literally matters of life-and-death. The most obvious examples are robotic systems such as self-driving vehicles or autonomous weapons systems, but disembodied systems can also make ethically meaningful, entirely autonomous decisions (e.g., loan approvals, medical diagnoses). For these (semi-)autonomous systems, we cannot assign ethical responsibility to the human user of the technology, precisely because there is no immediate human user. Of course, we could try to save the neutrality thesis by ensuring that there is a human involved in these decisions (though that move risks converting the human into a "moral crumple zone" who exists solely to bear moral responsibility; Elish, 2019). If a human were always involved, then we might hope that ethical analysis could remain focused on uses, rather than the technology itself. Unfortunately, this hope is empirically implausible: there are too many contexts and uses that require (semi-)autonomous digital technologies, so we must engage with the ethics of the decisions made by such systems. Unsurprisingly, there is a large (and rapidly growing) literature about ethical analyses, challenges, and principles for behaviors of autonomous systems (e.g., Anderson & Anderson, 2015; Lin, 2016; Millar et al., 2017; Sparrow, 2016; or many papers at the AI, Ethics, & Society conferences). At the same time, though, autonomous systems are still relatively uncommon; most digital technologies involve human decisions at various points. We might thereby hope that the neutrality thesis could be mostly saved, at least if the sphere of autonomous decisions is sufficiently small. Unfortunately, there is a challenge that is discussed much less frequently, but impacts all digital technologies, not just autonomous systems.

The second problem with the neutrality thesis for digital technologies is that the creation of digital technologies involves an enormous number of ethical choices that thereby embed or implement values in the technology itself. Most obviously, digital technologies are almost always designed or optimized for success at a specific task, but 'success' is an ethically substantive term. For example, consider a loan approval algorithm: should it be optimized to maximize societal benefit through credit, or profit for the lender, or empowerment of underserved communities, or some other outcome? We can readily develop algorithms for any of these goals (assuming that we have appropriate data and measures), but almost certainly, we cannot develop an algorithm that maximizes all of these goals simultaneously. That is, the decision about what to optimize is not a technical one, but rather is an ethical one about which problems are more important. And by making one decision rather than another, we have produced a digital technology that prioritizes one ethical value rather than another, regardless of how the algorithm is actually used in the future. This particular example shows how ethical choices during technology development can imbue digital technologies with values, but the challenge for the neutrality thesis arises throughout the technology "pipeline."

There are many different ways of describing how we move from idea to digital technology. For simplicity, a high-level caricature of the technology "pipeline" is provided below. Importantly, nothing significant depends on this particular caricature; any other framework for understanding technology development would lead to the same conclusions about the role of ethics. And although this caricature (including the language of a 'pipeline') might suggest a unidirectional flow from one stage to the next, matters are rarely so simple when building a digital technology. Invariably, insights and developments at a later stage will require us to revisit decisions and choices at earlier stages. With those caveats in mind, consider these six stages in the technology pipeline:

1. *Identify*: What is the problem (or problems) to be solved, and in what contexts, with this digital technology?
2. *Design*: What constraints—technological (including data), financial, legal, regulatory, performance, societal, ethical—do we face in developing this technology? How strong is each constraint, and how do they interact with one another?
3. *Develop*: What digital technology best satisfies these various constraints?
4. *Deploy*: Who has access to this technology (including where and when)?
5. *Use*: How is the digital technology actually employed (and by whom) to solve the problem(s) in real-world contexts?
6. *Revise*: In light of what we have learned, how should the digital technology be adjusted (including potential changes to the contexts of use)?

These stages are obviously not perfectly separable, and as noted above, technology creation rarely moves through them in a unidirectional manner. Nonetheless, they provide a useful framework for recognizing the ubiquity of ethical questions throughout technology creation, including the ways that answers to those question imbue the technology with values. While a complete list of such questions would be far too long to be useful, a briefer survey can help to demonstrate the ubiquity of ethical choices and decisions in digital technology creation.

In the Identify stage, the focus is on the problem(s) that we are addressing with our digital technology. The question "What problems are we trying to solve?" is necessarily an ethical one, as the decision to address problem *A* (rather than problem *B*) implies a value judgment that *A* is more important than *B*. In addition, we need to ask whether a solution to problem *A* would potentially create new problems or challenges. If so, then we need to again make an ethical decision about which problems (including the potential new ones) are most critical. Importantly, these questions and decisions cannot be avoided; we cannot build technology without identifying its intended functionality (including problems that it will address), and that identification necessarily involves values, interests, and other ethical commitments.

The Design phase focuses on the constraints that we face in technology creation, as well as the creative development of design solutions that address those constraints in a satisfactory manner. Ethics and values play key roles in the articulation of a set of design constraints. For example, if we include financial constraints, then we are making the ethical decision that monetary value is important to this technology. The resolution of potential tradeoffs or conflicts between design constraints provides another set of ethical questions. If we decide, for example, that a certain level of performance is less important than avoiding intrusive data collection, then we have thereby made ethical decisions about the relative importance of functional versus privacy constraints.

The Develop stage is typically the primary focus of technology creators, as this stage involves the actual implementation of our digital technology. For example, if we are trying to solve a problem using machine learning, then this stage is when we actually find the best-fitting parameters for the algorithm or model, given our available data. As a result, this phase might seem to be purely technical; the problems, constraints, and designs from the previous two stages required ethical choices, but perhaps development can be the value-free implementation of those designs. However, ethical decisions arise even here. There will typically be multiple ways to satisfy our various constraints, and so development requires us to choose between them. Our constraints rarely specify absolutely every element of the technology, and the remaining decisions can involve an ethical dimension. As a concrete example, our constraints might not dictate a particular color of a button in an interface, but the choice of red versus green could be meaningful in terms of setting an expectation that pressing the button would stop versus start something.

In the Deploy stage, the questions and challenges are principally about access, which inevitably has a significant values component. We are rarely able to provide equal access to the digital technology to all individuals, particularly if we broaden our understanding of 'access' to include full, appropriate access to all capabilities of the technology. But if deployment benefits only some individuals or groups (or benefits them more than others), then we are thereby making a value judgment that some people have greater need for the technology than others. This ethical decision is particularly important to recognize when deployment is market-based. Those individuals who are able to pay do not necessarily have the most ethically important need, yet market-based deployment prioritizes those individuals over others.

Given access to some digital technology, we must then examine how Use occurs with it. Outcomes are a critical part of ethical evaluation of some technology; at the very least, we need to determine whether the technology actually helps to solve the intended problem(s) in the real world. One also needs to ask about potentially unethical variations or differences that result from use, as different contexts or knowledge can lead to radically different outcomes for people with access to the same digital technology. The Use stage also raises questions of oversight and monitoring, particularly for relatively new or under-tested technologies. In all of these cases, values, interests, and other ethical considerations should play a key role in our real-world decisions about how to translate a digital technology from the lab to the real world.

Finally, the development and use of any new technology will inevitably result in problems or mistakes, and the Revise stage provides an opportunity to address those issues. However, that opportunity requires answers to questions about which errors should be fixed, how they should be fixed, whose needs or problems remain unsolved by this technology, and so forth. Each of those questions requires ethical commitments that further shape the digital technology itself.

At every stage of technology creation, we are thus forced to confront ethical issues and questions. We might have hoped that new technologies would not require us to prioritize certain values or resolve conflicts between interests, but that hope is always in vain. Ethical questions are simply ubiquitous in technology creation. We should not pretend that technology could be value-neutral, but instead embrace the ethical questions by explicitly and openly asking them. The challenge then becomes: how do we appropriately *answer* these questions? We turn now to this issue of translating ethical considerations into practical guidance for technology creators, developers, users, regulators, and more.

TRANSLATING ETHICS

There are two intuitively plausible types of strategies to develop practices for ethical creation and use of digital technology. Unfortunately, only one of them is likely to succeed, but it is informative to first consider the problems with the other. In particular, the previous section described the many ethical and value-centric questions that arise at every stage of the technology creation pipeline. At the same time, moral philosophers have developed many normative ethical theories over the past centuries, so we might hope that we could simply translate those theories into practices. For example, we might hope that questions of tradeoffs could be directly translated into the language of our preferred normative ethical theory, and thereby answered by technology creators. Or we could change deployment practices to include explicit, formal evaluation of the ethical permissibility of various possible strategies.

This hope has been most prominently expressed with regards to autonomous technologies such as robots, where many people have hoped that normative ethical theories could be literally written into the system's code (e.g., Arkin et al., 2011; and many others). One might hope, for instance, that a self-driving car could be coded to make explicit consequentialist calculations whenever the system must make an ethical choice, or to decide in accordance with a set of deontological principles. Unfortunately, this hope will not be feasible for most autonomous technologies, as they do not understand the world in the ways that we do. A self-driving car, for example, almost certainly does not use the same concepts as we do in our normative ethical theories. They are simply not programmed in ways that enable the explicit construction or implementation of some normative ethical theory. Thankfully, though, we do not actually need to be able to explicitly code ethics into an autonomous system. If the *human* creators—designers, developers, deployers, users—make ethical decisions in the creation of the autonomous technology, then the resulting system should itself make ethically defensible decisions (all else being equal). So could normative ethical theories play a central role in the human decision-making?

Unfortunately, there are two reasons to doubt the usefulness of explicit, conscious application of normative ethical theories, at least most of the theories that have historically been developed by ethicists and moral philosophers. First, digital technology creation almost always involves substantive values such as my personal interest in connecting with others (e.g., via social networking systems). In contrast, most philosophical work on normative ethics has focused on very high-level and universally applicable values and interests, such as universal human rights. And even when normative ethical theorizing has engaged with more substantive values, it has usually left open the exact weighting or tradeoffs between those values. As such, the normative ethical theories will not provide much guidance to human technology creators. Second, as outlined in the previous section, the technology creation pipeline is quite complex, involving many stages and multiple feedback loops. There are thus many potential places of intervention and decision-making. In practice, there will rarely, if ever, be a single normative ethical theory that is appropriate for all such decisions. The complexity of technology creation precludes the possibility of coordination on a single, universal ethical theory to guide all relevant decisions. We cannot explicitly, consciously use normative ethical theories to make (ethically) good decisions throughout technology creation.

We must find a different way to implement or operationalize our values into good decisions. More specifically, we need a "translational science" of digital ethics that discovers better and best practices to support and enable ethical decision-making at all stages of technology creation. This type of translational ethics is not the simplistic application of existing "basic" research to specific problems. The ethical challenges in technology creation involve value-, situation-, and technology-specific constraints that must be incorporated into our practices. Rather than starting with high-level normative ethical theories, digital ethics in the real world requires the multidisciplinary integration of insights from research in ethics, cognitive science, organizational behavior, sociology, legal studies, and much more. Conversely, we cannot stop with the articulation of high-level "principles for ethical AI," but instead must convert them into real-world, practical guidance. We must draw from a range of different disciplines to develop practices at each pipeline stage that make good decisions more likely. For convenience, we can think about these changes as falling into four broad (and overlapping) categories: people, processes, policies, and partnerships.

First, we need to think about the *People* who are making these varied decisions. In practice, these individuals rarely have training in ethical decision-making. Algorithm developers, for example, usually have a background in computer science or statistics, while technology regulators typically come from policy backgrounds. When ethical questions are raised during technology creation, a common reply is "we know how to build tech, not answer those kinds of questions."[1] Of course, difficulty answering a question does not thereby make the question irrelevant. Ethically impactful decisions are still being made throughout technology creation, albeit implicitly in the decisions to pursue one technological option rather than another. We thus must ensure that the people in the technology pipeline are appropriately trained and empowered.

One pathway to support people is through explicit education of those individuals who are already part of the technology creation pipeline. Many organizations have begun to produce educational materials around broad topics such as "Responsible AI" as well as focused topics such as "Obtaining user consent." Some of these educational materials are suitable for a range of sectors, while others are highly sector-specific (e.g., for highly regulated sectors such as finance or healthcare). These products will undoubtedly continue to grow and improve over time. At the same time, we should recognize the potential limitations of explicit instruction in ethical reasoning and decision-making. There has been a large amount of pedagogical research on teaching engineering ethics and bioethics in a range of contexts. That work has shown that standalone courses can have a positive impact on subsequent ethical reasoning and decision-making, but their impact seems to be notably less than when ethical issues are taught as part of "normal practice" (e.g., Davis, 1993, 2006; Corple et al., 2020; though more studies are needed, see Hess & Fore, 2018). Ethics-specific educational materials developed specifically for technology creators are likely to be less effective than integration of ethical considerations throughout their original technology-centric training.[2]

A different pathway to support people is through expansion of teams to include individuals with appropriate ethical (and other) training. Nowadays, user experience (UX) designers are a completely standard part of the development team for a new digital technology, but this was not always the case. Practices have shifted over the past few decades so that any serious development team will have access to people with specialized UX design skills and training. The same shift has not happened for ethics (though some organizations have started down this path), but we can envision a future in which standard practice is to ensure that any technology creation team has access to people with appropriate ethical, psychological, and sociological training to help answer the value-centric questions that are ubiquitous in technology creation. Of course, the viability of this pathway depends on a supply of people with ethical training who also understand the processes of technology creation. Thankfully, many universities are actively developing and deploying educational programs that provide exactly this type of training. Significant open questions remain about how best to integrate these individuals into technology creation teams, but those issues are being actively addressed by researchers in disciplines such as organizational design and industrial psychology.

Second, we must translate ethical considerations into new *Processes* for digital technology creation. In many cases, unethical digital technologies—more precisely, technologies that fail to implement or realize the values that we want—occur because of relatively simple, avoidable errors and decisions in various stages of the pipeline. For example, suppose a developer creates a system that optimizes prediction of student dropout (at a university), but then the user of that system (elsewhere in the university administration) believes that the system predicts which students will have low grades. While dropout and poor scores are correlated, they are obviously not identical. And as a result of this (avoidable) miscommunication between the developer and user, the algorithm could be used in ways that are systematically biased against particular communities (Fazelpour & Danks, 2021). Or consider a decision to measure the performance of a social network platform in terms of mean user engagement rather than median user engagement. This seemingly technical decision can significantly change the performance of the platform, potentially towards being more unethical (i.e., failing to support people's values and interests). In practice, these decisions are often made by a relatively low-level employee who is thinking only about technical considerations, not ethical or societal ones, and so our processes need to shift.

As these two examples suggest, miscommunication or lack of awareness within our pipeline processes can often lead towards unethical technology. In many cases, we do not necessarily need to adjust our processes in deep ways as much as we need to ensure that all of the relevant actors understand the nature and implications of the decisions that they are making. Consider a decision such as frequency of querying a server to find out if anything has changed. This decision might seem unrelated to any ethical concerns, but that will depend on the particular domain of use and application. We cannot know in isolation whether that query frequency will support people's values and interests; knowledge of the role of that choice in the larger system is required. In general, many "standard practices" in digital technology development—modularization of code, changing variable names to abstract codes, etc.—may allow for increased efficiency, but do so at the cost of reduced understanding by the decision-makers (again, perhaps relatively low-level individuals).

A natural worry about these arguments is that they risk inducing paralysis within the technology creation pipeline. We obviously cannot spend two weeks analyzing every little decision, just in case it might happen to have some potential (no matter how small) for ethical impact. Rather, we need to find ways to adjust our processes so that they appropriately balance our many different values and interests, including ones like "need to ship our product soon." Thankfully, these kinds of process adjustments are quickly being developed, and many are available for implementation.

One example focuses on a persistent challenge in technology creation: data collection often proceeds separately from data analysis, and similarly for model creation and model use. Different skills are required for the different tasks, and division of cognitive labor is important for organizations. Moreover, there may even be legal barriers to having multiple steps done by the same person, or open communication between the individuals working on each step. Datasheets (Gebru et al., 2018) and model cards (Mitchell et al., 2019) are two process innovations that aim to reduce or eliminate problems that can arise from this kind of distributed work. Both instruments were developed to encode the key information that will be needed for someone else in the pipeline, but without requiring full access to, or disclosures about, the component itself. With a model card, for example, I can know the contexts in which this model is likely to be useful, even if I do not know anything about the inner workings of the model. And while it does take slightly more time to complete a datasheet or model card, the information is typically already known and available to the person completing the card. No additional research or investigation is required, just a few minutes to write up what was collected or analyzed.

A different kind of process adjustment focuses directly on identification of perhaps-unnoticed decisions. If people are unaware of the potential ethical and societal ramifications of a particular choice, then they could easily contribute to problematic digital technologies from ignorance rather than malice. One natural place for this type of problem is in the Identify stage of the pipeline, where one decides that a particular problem is worth addressing with technology. Identification is often done by a domain expert (perhaps in collaboration with a technologist or data scientist) who may not have any particular training in how to recognize or analyze potential ethical impacts (particularly since those impacts might be very removed from the current context). "Ethical triage" processes (Montague et al., 2021) can enable such individuals to quickly determine whether they should talk to someone with ethical (or social scientific) training before starting to Design or Develop. And of course, similar such tools could readily be developed for other stages in the pipeline. These kinds of process adjustments do not tell the person what they ought to do for the technology creation, but instead help them to know when they need to collaborate with others to bend the pipeline towards creation of ethical digital technologies.

Third, our *Policies* must be reconsidered by every actor in this sphere, including technology creation companies and regulators (whether government or private). While our focus in this chapter has largely been on individual- or perhaps company-level decisions, digital technology creation occurs within larger societal structures. Laws, regulations, internal company policies, industry-specific standards, international agreements, and much more all fall into the broad category of policies that can shape the constraints and decisions that are made during technology creation. Moreover, most of these fall under the scope of 'ethics' as outlined earlier, though the values, decisions, and interests are at the group level rather than individual. If we were to ban certain digital technologies, for example, then we presumably would avoid creating unethical versions that fail to support people's values.[3] Policies usually only change slowly, and so they are sometimes thought to be a poor mechanism to influence technology creation. However, exactly because they change slowly, policies can provide a stable context for creation: all of the relevant actors can assume that the relevant policies probably will not change significantly over the relevant timeframes. As a result, policies can have a far-reaching and long-lasting impact.

One core challenge, however, is that current policies often are a poor fit for digital technologies, for at least two different reasons. First, many policies about technologies were designed for systems that cannot be easily duplicated or replicated, such as buildings or airplanes. The policies assume that there are a limited number of items to be governed, and that there will be time or financial costs to significantly increasing that number. In contrast, digital technologies—particularly those that are "disembodied" such as software or algorithms—can usually be duplicated in a low-cost or zero-cost manner. The creation challenge is to build the first instance, not to create additional instances. As a result, policies about, for example, surveillance systems do not necessarily apply to online spaces since those governance rules assume that there are pre-existing barriers to surveillance expansion, whether physical, logistical, or financial. Online surveillance is extremely low-cost by comparison, particularly when filtered by other digital systems. As a result, the relevant power and value relationships can shift radically compared to what is assumed by the policy. Many current political efforts, including the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), aim to rebalance these power relations.

Second, many policies assume (perhaps implicitly) that the system will be deployed only in known environments. For example, most automotive standards for safe operation assume that the vehicle is not being driven underwater. As this example shows, the relevant context might be quite broad, but it is nonetheless assumed to be known. In contrast, many digital technologies are developed without meaningful, known constraints on the contexts of deployment and use. In some cases, this lack of constraint is an important design feature for the technology, as it enables it to be (in theory) used in arbitrary new environments. In other cases, this lack of knowledge is actually the reason to look to a digital technology, as when autonomous technologies are developed to make intelligent decisions in unforeseen situations. In either case, we need to consider different kinds of regulatory and policy frameworks since much less is known ahead of time. For example, we can draw inspiration from regulation of pharmaceutical interventions (e.g., drugs), where regulators typically use a dynamic, staged, regulatory process precisely so that they can learn and adapt the use of the pharmaceutical as real-world feedback is obtained. While some adjustments need to be made for digital technologies, the case of pharmaceuticals provides an important proof-of-concept, as well as an analogue, for dynamic regulation of digital technologies to help ensure that they support people's values (London & Danks, 2018). We can translate our ethical needs and interests into society-level changes, not just local practices.

Fourth and finally, the translation of ethical considerations into practice must acknowledge the fact that technology is no longer created by a single individual or company, but rather involves tightly connected *Partnerships* between many different organizations. And just as our Processes must ensure appropriate awareness and communication during digital technology creation, inter-organizational relationships must similarly be adjusted so that the resulting technology use appropriately supports people's values. Moreover, the challenges of awareness and communication are even harder to overcome when working across organizational or institutional boundaries. While different parts of a company might sometimes fail to talk with one another, they nonetheless are (usually) permitted to do so. In many cases, though, the relevant people at different companies are contractually forbidden to speak to one another about relevant details of the digital technology. Hence, there need to be formal mechanisms that enable transfer of key information across these boundaries while preserving intellectual property and trade secrets.

This challenge is now manifesting in a particular issue that we can call *ethical interoperability*. Different organizations are increasingly developing, promulgating, and using particular ethical principles for their digital technology. Of course, as we noted above, these principles need to be translated into practices, but they nonetheless can have significant impact on an organization. For example, the United States Department of Defense has endorsed a principle that all its AI must be "Equitable," thereby incurring an obligation to develop and deploy AI technology in particular ways (and not others). The Partnership challenge of ethical interoperability is: How can we reconcile different sets of ethical technology (use) principles? If organization A develops digital technology that satisfies its principles, then when can organization B use that same technology despite B's different principles? There are many cases where the same technology will satisfy multiple sets of ethical (use) principles, so we need some way to assess ethical interoperability. And it would be natural to try to address ethical interoperability by close examination of the practices to see if they might satisfy multiple, different principles. Currently, however, such assessment methods are essentially completely unknown.

FUTURE RESEARCH DIRECTIONS

This chapter has aimed to provide a framework for thinking about digital ethics as a practical endeavor, rather than a purely theoretical one. This framework also enables us to quickly recognize the many open questions and challenges that we face. Digital ethics can only become a reality if we provide robust, validated practices and interventions for all stages of the technology development pipeline, and across all four Ps (people, processes, policies, and partnerships). Some examples of better practices were provided in the previous section, but those are merely the start. There are many gaps in our knowledge about how to have ethical digital technologies. As just one example, we do not currently have effective, low-cost ways of training people about ethical revision of digital technology. Such training could surely be developed and empirically validated, but that will require future research. More generally, we lack implemented and tested solutions for almost all stage-category combination; one goal for this chapter is to essentially lay out a roadmap for research questions to be asked, all of the form "How can CATEGORY be changed so that practices in STAGE better support people's values and interests?"

Relatedly, we need to work that we identify the best "owner" for these various innovations in practice. Some are best handled by corporate executives, or by developers working directly with data, or by end-users, or by policy-makers, or by some other group entirely. Very few best practices apply universally to all stakeholders, so we need to ensure that our changes are targeted at the right individuals. Many of these "owners" will be defined by their goals and interests, not necessarily their institutional roles, and so we should take additional care to describe the owners in ways that do not presuppose a particular title. For example, a small non-profit organization might not have a "Chief Data Officer," but nonetheless should have someone who focuses on privacy and other concerns about the data that they collect. Or consider the ethics review committees that are increasingly appearing in technology companies (e.g., Microsoft's Aether Committee): these groups can provide important guidance, but we should not assume that every digital technology developer will have access to such a committee. And of course, some of these innovations are likely to have larger impact than others, so we should work to identify those that are more important in the short-term versus longer-term priorities.

As we collectively move towards better practices, we should also aim to be pluralists about our approaches and efforts. Almost certainly, there will not be any single discipline or approach that can solve all of our research challenges. We will need to draw from a range of disciplinary perspectives and methods while still insisting on rigor and clarity. For example, some questions may require the tools of social psychology, while others might draw on methods from analytic philosophy. We should be open-minded about our approaches to digital ethics, aiming to use whatever methods, concepts, and frameworks are appropriate. Digital ethics is necessarily and inevitably a collaborative effort that depends on a range of experiences, disciplines, paradigms, and approaches. As a result, we should strive (as a community) towards a "big tent" attitude that recognizes that contributions and advances can come from many different sources. Diversity of all types will be absolutely critical as we strive towards a robust, applicable digital ethics.

CONCLUSION

The language of 'digital ethics', along with calls for its importance, has become increasingly present in discussions about our digital technologies. There is a growing awareness that *something* is wrong with the ways that we design, develop, and deploy those technologies. And while the exact diagnosis of the problems differs between thinkers, they largely agree that ethics must, in some way, be an important part of the solution. Of course, the exact form and content of that solution is often under-specified, or articulated only at the highly abstract level of ethical principles. A meaningful digital ethics that actually leads to technology that better supports our values involves the conversion of these principles into useful, tangible practices. However, this conversion requires substantive research to translate the abstract insights into useful practices. Just as translational medicine converts biomedical research into clinically useful guidance, we need a translational digital ethics to yield useful changes to our current technology creation pipeline.

This vision of digital ethics is quite different from the most common paradigms currently used in technology creation. Most notably, the vision presented here starts with the recognition that ethical decisions are made throughout technology creation, so ethical analyses cannot be treated as optional or as one last checkbox on the way to deployment. The standard practices for many present-day technologists involve an exclusive focus on the (seemingly) “purely technical” problems to create the technology, followed by a consideration of the ways that it might go wrong. Guardrails and guidelines are added only at the end, and only if there is appropriate time to reflect on the potential ethical and societal impacts. In this paradigm, ethical analyses are comparable to commenting your code: it’s a nice thing to do if you have time, but it happens only at the end and very rarely changes anything important. In contrast, we need a digital ethics that is fully integrated into the practices of digital technology creation. A better parallel would thus be user interfaces: everyone now realizes that interfaces are built for essentially all software (even if you wish that you did not have to build one), and so we should employ our best science and theories to build good and usable interfaces. Similarly, digital ethics must be incorporated throughout all technology creation practices.

This type of translational ethics is still quite new in the digital technology space. We currently lack answers to many critical questions, such as how to identify stakeholders in a principled manner or how to determine whether a detailed ethical analysis is required. Thankfully, these kinds of questions are now being asked, and so the translational (digital) ethics is starting to be built. This research will inevitably require inter- and multi-disciplinary collaborations, precisely because digital technologies are, like (ethical) values, now ubiquitous and influential in many areas of our lives. Technologists do not need to know all of the answers, as that would require a skillset spanning many disciplines. But they do need to be able to collaborate closely with people from other disciplines to translate ethical insights into actionable, practical changes in our people, processes, policies, and partnerships.

ACKNOWLEDGMENT

Significant parts of this chapter were written while the author was Professor of Philosophy & Psychology at Carnegie Mellon University.

REFERENCES

Anderson, M., & Anderson, S. L. (2015). Toward ensuring ethical behavior from autonomous systems: A case-supported principle-based paradigm. *Industrial Robot: An International Journal*, *42*(4), 324–331. doi:10.1108/IR-12-2014-0434

Arkin, R. C., Ulam, P., & Wagner, A. R. (2011). Moral decision making in autonomous systems: Enforcement, moral emotions, dignity, trust, and deception. *Proceedings of the IEEE*, *100*(3), 571–589. doi:10.1109/JPROC.2011.2173265

Corple, D. J., Zoltowski, C. B., Kenny Feister, M., & Buzzanell, P. M. (2020). Understanding ethical decision-making in design. *Journal of Engineering Education*, *109*(2), 262–280. doi:10.1002/jee.20312

Davis, M. (1993). Ethics across the curriculum: Teaching professional responsibility in technical courses. *Teaching Philosophy*, *16*(3), 205–235. doi:10.5840/teachphil199316344

Davis, M. (2006). Integrating ethics into technical courses: Micro-insertion. *Science and Engineering Ethics*, *12*(4), 717–730. doi:10.100711948-006-0066-z PMID:17199146

Dotan, R. (2020). Theory choice, non-epistemic values, and machine learning. *Synthese*. Advance online publication. doi:10.100711229-020-02773-2

Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, *5*, 40–60. doi:10.17351/ests2019.260

Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, *16*(8). Advance online publication. doi:10.1111/phc3.12760

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press. doi:10.7551/mitpress/7585.001.0001

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2018). *Datasheets for datasets*. https://arxiv.org/abs/1803.09010v7

Hess, J. L., & Fore, G. (2018). A systematic literature review of US engineering ethics interventions. *Science and Engineering Ethics*, *24*(2), 551–583. doi:10.100711948-017-9910-6 PMID:28401510

Lin, P. (2016). Why ethics matters for autonomous cars. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous driving: Technical, legal and social aspects* (pp. 69–85). Springer. doi:10.1007/978-3-662-48847-8_4

London, A. J., & Danks, D. (2018). Regulating autonomous vehicles: A policy proposal. In *Proceedings of the 2018 AAAI/ACM Conference on artificial intelligence, ethics, and society* (pp. 216-221). Association for Computing Machinery. 10.1145/3278721.3278763

Millar, J., Lin, P., Abney, K., & Bekey, G. (2017). Ethics settings for autonomous vehicles. In L. Patrick, R. Jenkins, & K. Abney (Eds.), *Robot ethics 2.0: From autonomous cars to artificial intelligence* (pp. 20–34). Oxford University Press.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the 2019 Conference on fairness, accountability, and transparency* (pp. 220-229). Association for Computing Machinery. 10.1145/3287560.3287596

Montague, E., Day, T. E., Barry, D., Brumm, M., McAdie, A., Cooper, A. B., Wignall, J., Erdman, S., Núñez, D., Diekema, D., & Danks, D. (2021). The case for information fiduciaries: The implementation of a data ethics checklist at Seattle Children's Hospital. *Journal of the American Medical Informatics Association: JAMIA*, *28*(3), 650–652. doi:10.1093/jamia/ocaa307 PMID:33404593

Simon, J. (2017). Value-sensitive design and responsible research and innovation. In S. O. Hansson (Ed.), *The ethics of technology* (pp. 219–236). Rowman & Littlefield International.

Sparrow, R. (2016). Robots and respect: Assessing the case against autonomous weapon systems. *Ethics & International Affairs*, *30*(1), 93–116. doi:10.1017/S0892679415000647

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

ADDITIONAL READING

Aizenberg, E., & van den Hoven, J. (2020). Designing for human rights in AI. *Big Data & Society*, *7*(2). Advance online publication. doi:10.1177/2053951720949566

Fjeld, J., & Nagy, A. (2020). *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*. Berkman Klein Center for Internet & Society. https://cyber.harvard.edu/publication/2020/principled-ai

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. doi:10.100711023-018-9482-5 PMID:30930541

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. doi:10.103842256-019-0088-2

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)* (pp. 59-68). New York: Association for Computing Machinery. 10.1145/3287560.3287598

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. doi:10.1016/j.clsr.2021.105567

KEY TERMS AND DEFINITIONS

Datasheets: A framework and tool for representing and encoding key features of data so that others can use those data responsibly and ethically.

Digital Technology: A product or artifact, perhaps non-physical, that manipulates digital representations to accomplish specific tasks.

Ethical Analyses: Systematic descriptions of the ethical risks, benefits, challenges, and opportunities of a particular technology.

Ethical Practices: Patterns of behavior that lead to more ethical decisions, actions, and outcomes.

Ethical Principles: Commitments or beliefs, often quite abstract, that guide ethical decision-making across a range of domains.

Model Cards: A framework and tool for representing and encoding key features of AI models so that others can use those models responsibly and ethically.

ENDNOTES

[1] A closely related reply is "we will implement ethics in our systems as soon as the ethicists tell us what to code." As the previous paragraphs showed, though, this reply is misguided in multiple ways.

[2] Of course, companies cannot travel back in time to change their employees' original education! Standalone ethics-centric training might be the best that they can do at the moment. Nonetheless, organizations should be aware of the limitations of this kind of training, and also the importance of hiring technology-centric individuals whose technology training involved a substantial ethical component.

[3] Of course, we also would miss the potential benefits—ethical and other—that the technology might provide.

Chapter 2
Beyond Tools and Procedures:
The Role of AI Fairness in Responsible Business Discourse

Ivana Bartoletti
Women Leading in AI, UK

Lucia Lucchini
Deloitte, UK

ABSTRACT

As artificial intelligence (AI) is increasingly being deployed in almost all aspects of our daily lives, the discourse around the pervasiveness of algorithmic tools and automated decision-making appears to be almost a trivial one. This chapter investigates limits and opportunities within existing debates and examines the rapidly evolving legal landscape and recent court cases. The authors suggest that a viable approach to fairness, which ultimately remains a choice that organizations have to make, could be rooted in a new measurable and accountable responsible business framework.

INTRODUCTION

As Artificial Intelligence (AI) is increasingly being deployed in almost all aspects of our daily lives, the discourse around the pervasiveness of algorithmic tools and automated decision-making appears to be almost a trivial one.

AI solutions are now driving the allocation of resources as well as shape the news and product items that individuals are exposed to: from credit scoring to facial recognition, predictive technologies to identify fraudsters with precision, youth crime prevention tools, algorithm-driven advertising… how far AI can go is already a reality we all live with daily.

The prolific number of cases showing an evident misuse of these technological tools and, often, the personal data within them, are paving the way to novel discussions around ethics and its role in the digital world. Whether through the scraping of the web in search of faces to be used as facial recognition training data (Hill, 2020), or through algorithms automatically rewarding private school pupils with

DOI: 10.4018/978-1-7998-8467-5.ch002

higher grades (Burgess, 2020), the politics of data (and of data classification) has become increasingly impossible to hide and ignore.

Recognizing this allows us – as a society – to question the role that ethics can play around the development, deployment, and use of technology. While public calls for regulatory scrutiny is on the rise and dominates headlines news, we are yet to define how agencies, governments, as well as private sector organizations can provide meaningful notice about an algorithmic decision-making output. This has led to the deployment of flawed automation, the consequences of which ultimately harm trust in technology and hinders human rights by limiting and/or locking individuals out of services and equal opportunities.

Over recent times, we have had litigation in courts (for instance with Deliveroo, Uber and others) that have successfully deliberated on the impacts that misuse of technology can have on individuals. However, as argued by Calo and Citron (2021), the limit with litigation is that it only addresses violations that are currently enshrined in the law. As further argued by Ajunwa (2019), “instead of focusing on novelty, we should focus on salience” (p. 1675), moving the conversation from technical solutions to a technical problem, to reformed public policy to root out and address issues arising from flawed automation.

This chapter’s premise is that the very essence of machine learning is to differentiate, which means that bias lies at the core of this technology. The bias we, as a society, should be mostly concerned about is the one that causes either allocational harm (not allocating a service or a good to someone unfairly) or representational harm (the perpetuation of inequality by, for example, encoding stereotypes in advertising tools).

To an extent, it can be said that the engineering of unfairness is the inevitable outcome of the politics of dataification and data collection, for example, basing decisions around offending using an existing database of offenders is likely to be unfair as the most vulnerable communities are the most surveilled. Therefore, it can be argued that for an organization to make a fair decision, the algorithm (and the data it is fed with) should be actively manipulated to produce a fair outcome. *Therefore, the question moves into why organizations would want to opt for fairness, if fairness is financially detrimental to their business.* Ultimately, a choice of fairness is inevitably a political, social, and ethical choice as opting for fairness may lead to less efficient outcomes for an organization.

For example, a company that sells house cleaning product, when using ML to identify potential consumers to target, it is likely to optimize its success by targeting those who more frequently by such products, which for historical reasons means women. Should the same company decide to be ‘fair’ and start promoting equally to men and women, the outcome would be more equitable, though possibly less financially viable as less people would click on the ads.

This paper argues that, as opting for fairness may not be the optimal financial solution for an organization, its formalization resides in the responsible business that is gaining ground, amid consumers’ demand for more equity and transparency.

However, without public accountability and methods to expose wrongdoings, the responsible business discourse will – like ethics - likely end up as a generic commitment and a public relation (PR) slogan. Therefore, this chapter argues that it is of paramount importance to bring the dimension of fairness as an active choice to the public scrutiny, and this requires both public awareness and accountability measures.

WHY IS AI DIFFERENT?

Technology has been with us forever, in a way or another. Not only it has been incorporated into our lives, but it has also driven and mediated the way we live. For examples, speed bumps on our road as a way to limit speed have been the tool used to form and shape our behavior and, as such, have acted as a proxy for the law. However, AI tools are of a different nature, not simply because of their ubiquity, but also because of the affordances they bear. Due to AI's "volume, velocity, and variety" (Gewirtz, 2018), AI holds the power to solidify, amplify, and perpetuate through automation existing inequalities, profoundly reshaping the way individuals commonly live and operate.

AI does not exist in isolation, rather, it permeates our daily lives under the full weight and applicability of existing laws, holding the potential for severe impacts to fundamental rights encompassing the right to privacy, dignity, freedom, and autonomy.

The transformative power of AI has long been discussed and debated. The European Union (EU) has now introduced a set of norms to govern AI's impact and behaviors surrounding its deployment, including the suggestion to ban certain applications and/or build a set of restrictions and constraints around AI systems that may cause harm. One key example is credit scoring, where denying the allocation of a loan may lead to substantial harm to an individual, affecting esteem as well as reducing individuals into real poverty.

The risk based approached the EU wants to enshrine in the law is in recognition of the affordances AI tools may have – meaning that the more power AI holds to lock people out of services; the more controls and accountability procedures need to be placed around AI to ensure they operate in an unaccountable way.

Furthermore, AI extends far more beyond its technological and social dimensions. Computational power relies on a few big tech companies providing the infrastructure and increasingly sharing the norms around AI, which has significant implications from the perspective of competition law. If we add that the geopolitics of undersea cables resembles that of oil and natural-gas pipelines, it is clear that a conversation around AI is a conversation around power and politics. Ultimately, it is for this very reason, that the role of ethics in this debate is of critical importance.

The Debate Around Fairness

Centuries of literature and philosophical thinking have revolved around the concept of fairness, leaving it still very much open to debate and deliberation. The issue surrounding fairness, especially in the context of data processing and the output of an AI system, lies in the complexity to define and establish a common approach that can ultimately unite and drive an organization. In the context of AI, fairness can be defined across a variety of disciplines.

From a data privacy standpoint, fairness relates to two elements. The first one is in relation to the *balance of power between the data controller and the user.* The second is in relation to the treatment of the more vulnerable, starting with children who require additional care and safeguarding. With regards to the EU General Data Protection Regulation (GDPR), AI systems using personal data are audited against that approach to fairness. However, privacy in the EU is seen through the lens of human rights, and this is not the same across the world. This is an important caveat, when looking at fairness in AI, insofar as it implicates an inherent association between fairness and human rights.

More recently through the Black Lives Matter movement, the topic of unfairness within technologies and the demand for equity has emerged vigorously. It is not by coincidence, that the most vocal criticism of an unfettered used of AI come from black academics and policy leaders, and that the most vocal and powerful criticism of algorithms used to govern, rule, decide, evaluate are indeed workers.

Returning to the example of credit worthiness, empirical studies examining the impact of algorithmic credit scoring on distributional fairness in consumer lending are instructive. In particular, a study based on US housing mortgage data points to the elimination of discrimination in loan origination, and a reduction in loan pricing discrimination against minority groups (here, Hispanic and African American borrowers) as a result of algorithmic credit scoring (although price discrimination nevertheless persists).

Similarly, a second study suggests that the use of algorithmic credit scoring could increase loan acceptance rates for Hispanic and African American borrowers.

However, the same minority groups were also more likely to receive higher interest rates with algorithmic credit scoring, and greater within-group dispersion of rates, as compared to White and Asian borrowers (Aggarwal, 2021).

The above is, in our view, quite emblematic of the situation we are in. *Without a clearly agreed and set out definition of fairness, the assumptions on an AI artefact can be misinterpreted.* For example, vulnerable consumers can be at risk from the use of behavioral insights derived by lenders through algorithmic credit scoring. Unscrupulous lenders can exploit data-driven insights to pursue more aggressive debt collection practices, targeted at the most vulnerable borrowers.

The Proliferation of Tools for Ethical AI

When GDPR came into force, it was perceived by organizations as an obstacle to developing infrastructure, forcing them to abandon contacts, relationships and advantages built over time. At the same time, GDPR forced organizations to innovate by rethinking solutions and finding new ways of operating in a compliant manner.

The introduction of the GDPR in Europe spurred a new wave of data privacy and data protection regulations. To this day, there are over 180 around the world that have recently introduced a data protection legislation. In the US, CCPA and CPRA have raised the bar in relation to consumer rights, and laws are emerging all across the country alongside the increased awareness of the inextricable link between *privacy* and *the policy of control* over individuals – a link that has emerged over recent months, from policing to work surveillance, and especially in relation to people of color.

It is likely that governance and regulation on AI will have the same impact. Introducing restrictions, both internal and external, will be perceived by businesses as a brake to innovation, as they impose the obligation to refocus on ethics. Simultaneously, however, this will also push organizations to think about how to reassess their innovation, how to adapt business models, and how to adopt decisions that ultimately hint the consumer as beneficiary of their products and services.

AI assurance / risk management tools are often evoked in relation to systems deployed, particularly by the public sector, as across the world inadequately resourced public sector entities and local authorities have turned to automation to reduce costs. Regardless of the numerous times these systems have been challenged for their opacity and lack of accountability, organizations have continued to adopt these tools, often by third party vendors.

Over the last few years, we have seen the emergence of a plethora of tools addressing AI pitfalls. These includes debiasing tools, tech mechanisms and procedures to ascertain whether a system is biased and adequately explainable, along with ways to address and remediate issues.

On the one hand, these tools, when adopted in isolation and without a codified and standardized approach to fairness, are often insufficient and may result in further damage. They can also be used as get out clause for removing wider social issues, including those identified at the core of the Black Lives Movement and/or those emerging from the pandemic, where the most vulnerable are disproportionately being impacted.

On the other hand, algorithmic tools can be practical enablers for organizations, and for the extended communities, to entertain a much wider discourse around the role of ethics. Furthermore, with the complexity of identifying discrimination, especially when it occurs in a subtle or by proxy way tools can be regarded as a useful instrument to help investigate discrimination beyond the traditional normative description (Wachter et al., 2021).

THE LIMITS OF TECHNICAL AND COMPUTATIONAL SOLUTIONS

As part of their AI development and deployment governance, organizations often deploy tools to identify how a model operate so that harmful bias can be identified and corrected. This may be particularly helpful amid the complexity and subtlety of how discrimination happens in AI system, which includes proxies.

Ex ante, organizations also establish discrimination-sensitive quality regimes for training data, and this may include accuracy, timelessness (that says that yesterday's data must not determine tomorrow's choices), completeness and representativeness.

However, the transformative nature of algorithms and the affordances they carry mean that a computational/mathematical solution alone is at risk of overlooking the social nature of technology. This is because a technological product is the outcome of a specific reality in which the product is conceptualized, developed, and deployed.

We note that there are two dichotomic approaches to tackle harmful bias. The first one says that we can respond to technological faults simply by using more technology. The second says that a technological fix is nothing more than the confirmation that the problem is instead sociological.

We argue that this dichotomic approach is not conducive to understanding the dual nature of a technological artefact. The EU AI Act provides an interesting approach to this as it follows the risk-based approach that is typical of the GDPR.

However, as mentioned by Floridi (2021), high risk AI could be an artefact that fails to work, as well as an artefact that, if put to work, can be high risk. For this reason, the risk-based approach could lead us to the risk of confusing and wrongly codified the most suitable definition of high risk. In other words, the narrative around risks cannot be focused only on the risk of operating in a way that is considered wrong, as it could overshadow the long-term consequences of a machine operating well.

THE CONCEPTUALISATION OF AI AS A SYSTEM TO CREATE HUMAN ACCOUNTABILITY FOR TECHNOLOGY

Inevitably, the EU AI Act, alongside other governance measures on the use of AI systems, may foster a culture of transparency. Court cases leveraging the GDPR to uphold individual rights also confirm that, although in a limited way, privacy law has a role to play.

The Council of Europe, much larger than the European Union, is also working on a (partially?) binding AI Treaty with a specific focus on human rights.

Regulating the use of AI systems is going to be complex – but even more complex is going to be the codification of how AI systems can actually help the common good. In other words, how will we be able to fully understand the dual nature of technology (social and technical) in a way that enables us to leverage the latter to improve the former?

In a personal reflection on the Ofqual A level 2020 algorithmic debacle, former Ofqual chair Roger Taylor (2020) makes a very interesting point that the mistakes were made by humans, not machines. He states: "the exam grades debacle of 2020 has been blamed on a malfunctioning algorithm. But by blaming the algorithm, we risk missing the most important lessons on mistakes that were made." And added: "The problem was not the algorithm; it was what we were trying to do with it: it was human decision making that failed" (p. 1).

In our view, this personal account is important as it highlights the risks of diluting the social and the technical, dispersing accountability for either of those. As further presented by Ben Green and Amba Kak (2021), "Rather than prompt a superficial "human-in-the-loop" policy fix, the material harms caused by A.I. must trigger a re-evaluation of whether many of these systems should be used at all and greater accountability for the real human (and institutional) decision-makers behind these harms" (para. 12).

For this reason, we argue for the conceptualisation of an AI artefact as a system, rather than a model. By codifying an AI artefact as a system, the social element of the artefact can be conceived at every step:

1. Objective setting
2. Data collection
3. Model development
4. Model testing
5. Model monitoring

Each step will need to be treated separately, encompassing both a technical and a social approach. Clear accountability around the methods used, a full system of assurance operating at each stage of the system, and clear audit logs of the decisions taken are essential elements for accountability. Additionally, an obligation to report the measures taken, as part of annual reporting, could provide further transparency and contestability.

Ultimately, we argue that the conceptualization of AI artefacts as socio-technical systems will lead to the systematization of AI as part of the responsible business agenda, holding organizations accountable for the choices they make, and for the clear explanation of why they choose a technological artefact to solve an often social and political problem.

FUTURE RESEARCH DIRECTIONS

As the role of ethics in new technologies expands to new remits, with AI fairness representing an inherent component, the authors of this Chapter believe that the interrelation between the human and the AI artefact in decision-making will increasingly require further investigation. While current data protection regulations, including the GDPR, provide a workable definition and rights surrounding automated decision-making, the intrinsic nature of automation in the decision-making process is constantly in development. This, along with the necessity to integrate an ethical and fair approach to AI implementation and deployment, is bound to change the way the human will be expected to interact.

CONCLUSION

This Chapter argued that fairness may not often be the optimal financial solution for an organization, as it requires the manipulation of the data and the AI artefact that processes it. For this reason, choosing to adopt fairness in algorithmic decision-making is an inherently political, social and ethical choice. Accordingly, the authors argued that AI artefacts should be conceptualized as socio-technical systems, encompassing both a technical and social approach.

To achieve this, the Chapter acknowledged the increased public calls for regulatory scrutiny around the use of technologies and the data behind them, along with the current inability to define how agencies, governments, as well as private sector organizations can meaningfully develop, deploy, use, and provide notice about an algorithmic decision-making output. Accordingly, the authors noted that the deployment of flawed automation will inevitably harm trust in technology, perpetuating existing inequalities at the risk of hindering the fundamental rights of privacy, human dignity, freedom, and autonomy.

While a plethora of tools attempting to address AI pitfalls are flooding the market, including debiasing activities and algorithmic auditing tools, the Chapter evaluated the ways in which such tools, when adopted in isolation and without a standardized approach to fairness, can lead to further damage by removing the wider socio-political dimension of algorithmic decision-making.

Ultimately, upcoming risk-based regulations, i.e., the EU AI Act, will force organizations to reassess their innovation, consequently requiring the adaption of current business models, decision-making mechanisms, and the way organizations will choose to adopt ethics.

REFERENCES

Aggarwal, N. (2021). The norms of algorithmic credit scoring. *The Cambridge Law Journal*, *80*(1), 42–73. doi:10.1017/S0008197321000015

Ajunwa, I. (2019). The paradox of automation as anti-bias intervention. *Cardozo Law Review*, *41*, 1671–1742. https://scholarship.law.unc.edu/faculty_publications/491

Burgess, M. (2020, August 20). *The lessons we all must learn from the A-levels algorithm debacle*. Wired. https://www.wired.co.uk/article/gcse-results-alevels-algorithm-explained

Calo, R., & Citron, D. K. (2021). The automated administrative state: A crisis of legitimacy. *Emory Law Journal*, *70*(4), 797–845. https://scholarlycommons.law.emory.edu/elj/vol70/iss4/1

Floridi, L. (2021). The European legislation on AI: A brief analysis of its philosophical approach. *Philosophy & Technology*, *34*(2), 215–222. doi:10.100713347-021-00460-9 PMID:34104628

Gewirtz, D. (2018, March 18). *Volume, velocity, and variety: Understanding the three V's of big data*. ZDNet. https://www.zdnet.com/article/volume-velocity-and-variety-understanding-the-three-vs-of-big-data

Green, B., & Kak, A. (2021, June 15). The false comfort of human oversight as an antidote to A.I. harm. *Slate*. https://slate.com/technology/2021/06/human-oversight-artificial-intelligence-laws.html?via=rss_socialflow_twitter

Hill, K. (2020, March 18). The secretive cimpany that might end privacy as we know it. *New York Times*. https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html

Taylor, R. (2020). *Is the algorithm working for us? Algorithms, qualifications and fairness*. Centre for Progressive Policy. https://www.progressive-policy.net/downloads/files/Is-the-algorithm-working-for-us_Roger-Taylor.pdf

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. Advance online publication. doi:10.1016/j.clsr.2021.105567

KEY TERMS AND DEFINITIONS

AI Ethics: The systemic conceptualization of 'right' and 'wrong' based on five key themes: beneficence, non-maleficence, autonomy, justice, explicability.

Algorithmic Bias: It refers to the unintended and potentially harmful skewing of algorithmic predictions.

Equality: The belief that all humans are fundamentally equal and deserve equal treatment.

Equity: The value that drives the reduction of avoidable inequalities between people in society.

Fairness: To be distinguished between the sociological and the mathematical meaning. From the sociological perspective, fairness defines the way some people are being treated in a society and it is heavily based on ethical values.

Justice: Adequate adherence to the standards established in a given society.

Section 2

Digital Ethics, Equity, and Human Rights

Chapter 3
Ethics, Digital Rights Management, and Cyber Security:
A Technical Insight of the Authorization Technologies in Digital Rights Management and the Need of Ethics

Ali Hussain
Faculty of Computer Science and Information Technology (FSKTM), Universiti Malaya, Malaysia

Miss Laiha Mat Kiah
Faculty of Computer Science and Information Technology (FSKTM), Universiti Malaya, Malaysia

ABSTRACT

Cloud content hosting and redistribution is enabling convenient and easy access to online content thereby accelerating the adoption and penetration of internet in past two decades. The current Industry 4.0 revolution and adoption and acceleration efforts are leveraging cloud computing as a means to store, retrieve, and share data. This makes the internet a relatively vulnerable to content abuse and increase the demand of clear consent before data consumption and redistribution. The growth of cloud computing and management technologies is penetrating in the market, and digital rights management (DRM) practices are needed for better and ethically safe online space. This chapter talks about state-of-the-art DRM paradigms being proposed in the literature and critically discusses their technical performance, flexibility, and immutability challenges. This chapter will clarify internet governance implementation roadmap for Industry 4.0 revolution by critically analyzing the cloud technology stack and ethical features by advocating Cloud DRM.

DOI: 10.4018/978-1-7998-8467-5.ch003

INTRODUCTION

The increasing popularity and use of the Internet, has resulted in huge amounts of personalized content being stored in Cloud storage. This private data is accessed and shared using heterogeneous devices such as mobile, web, and IoT (Salman et al., 2018). The open and inclusive nature of the Internet allows content to be redistributed, without explicit consent, which may infringe Intellectual Property rights. With the recent growth in the utilization of cost-effective and easy-to-scale Cloud computing, it has become a popular way of storing online content. However, the protocol and mechanisms in Cloud computing need to be questioned for the greater safety and security of the content they store. This is because of a paradigm shift from client servers to the distributed nature of Cloud protocols, which bring new challenges for content security. Digital Rights Management (DRM) is a technology used to safeguard data over the application layer from illicit usage, using some form of encryption. DRM technologies for Cloud content need to be explored to protect Cloud data from unauthorized usage and efficient revocation. (Q. Huang et al., 2013; Koulouzis et al., 2018; Xu et al., 2017). The unauthorized Cloud content sharing or maneuvering access licenses may infringe copyrights and online content security policies.

It is necessary to apply technical solutions to mitigate this content misuse threat to online safety. However, implementing strong and computation exhaustive encryption in a distributed mode, as well as content serving, introduces many performance bottlenecks and constraints such as synchronization, latency, low processing and storage, and client and communication overheads (Das et al., 2015; Ma, Jiang, et al., 2018). This overhead needs to be reduced by researching robust encryption in DRM solutions to meet the practical and flexible need for content protection in Cloud computing.

Gartner forecasts that 6.4 billion connected things will be in use worldwide in 2016, up 30 percent from 2015, and increasing to 20.8 billion by 2020. A large percentage of these devices are going to be Bring your own device (BYOD), which carries a combination of personal and business data, and has major vulnerabilities – putting your enterprise data at risk. Attackers have shifted their focus from secure enterprise environments to these vulnerable devices, which gives them an easy gateway to enterprise data. Mobile Device Management has become an important risk mitigation measure for every organization.

Encryption of data at rest will evolve to target the resources, rather than where the resources are stored. Applications have become more complex, with some of them running on-premise and some in the Cloud. There is a business need for data to be transferred across these applications and thus blindly encrypting the entire repository where the content is stored will be rendered ineffective. With the amount of data handled increasing every day, it is too expensive and unmanageable for organizations to encrypt data blindly. Companies will have to start moving to solutions that can dynamically protect information based on the criticality of the data to comply with business and regulatory requirements.

Based on a survey conducted by Gartner, IT security ranks second in the list of priorities for corporate investment. This is primarily due to the increase in data breaches in the recent past and the impact these incidents have on the companies' credibility.

So how can organizations secure their data effectively? This will require a multi-pronged approach to plug all potential loopholes and safeguard the 'crown jewels' of the organization. The first step in this approach is to have well-defined security policies that prescribe what information needs to be safeguarded and how it needs to be protected. Companies need to take a top-down approach to this, with the policies percolating from the CEO down to the lowest levels of the organization. A key factor is to ensure that the policies remain relevant to changing business needs and can be changed quickly to reflect new requirements. Solutions should automatically apply protection to reduce human error. Employee awareness about the latest cyber threats and protection methods is also important to ensuring security is properly enforced.

The constantly evolving data security risks are a huge challenge for the security teams, but with the right policies and tools, they can secure their most critical information. A data-centric security strategy that is flexible, dynamic, and provides central visibility needs to be a key part of any security portfolio.

This chapter will talk about state-of-the-art DRM paradigms being proposed in the literature and critically discusses their technical performance, flexibility, and immutability challenges. It will also contribute to help streaming an Internet governance implementation roadmap for the industry 4.0 revolution by critically analyzing the Cloud technology stack and ethical features and outlining what is needed for Cloud DRM.

BACKGROUND

In the recent past, the growing access to online digital media has been the focus of research and has gained significant importance to serve the content, meaning that content protection with optimal performance and flexibility is a significant challenge. However, the volume of digital data consumed by end-users is also increasing tremendously. This is due to the rapid improvement in the networking infrastructure, that boosts the usage of information technologies such as audio, video, social media, and Cloud storage as collaboration platforms. This increasing popularity and use bring copyright and data security changes. This is because the traditional security mechanisms, at application layers, are not directly applicable for an easy DRM implementation in the Cloud. Therefore, the content security mechanisms such as DRM methods are necessary for the safety and security of Internet content in protecting user's digital rights such as intellectual property, data confidentiality, data integrity, and preventing unauthorized data access, which should meet the practical application requirements such as lightweight and flexibility.

The growing use of DRM is evident in its global adoption and market share. In 2012, the Global DRM market volume was about $1,290 million USD and is predicted to increase to $2,024 million USD by 2024. The DRM 's increasing market volume is due to its increasing use in software, automation, and security layers, being implemented in office and offshore collaboration storage, and by the public, private as well as industrial-scale applications. The benefits of DRM need to be leveraged in the modern applications of Cloud computing, for content storage for large volumes of multimedia files and documents. In the Cloud, extending DRM features requires more stringent technology, to protect digital assets such as e-books, music, movies, etc. Heterogeneous devices accessing the Cloud content suffer poor quality of service due to many protocol bottlenecks. Modern Cloud storage such as Google Drive, OneDrive, Dropbox, and Box are commonly being used as content storage. These Cloud technologies and smart devices are being used for storing and sharing user's online digital content (H. Lee et al., 2016b). DRM in the Cloud will allow content owners to share the content with different users, and allow consumers to access the data according to the access policy set by the content owner.

This delegation of access in Cloud computing is growing in importance due to its many benefits including flexibility and safety. DRM is applied to safeguard the Cloud data as an extra wall of defense from security breaches. The encryption in DRM helps to save unencrypted data exposure and limits its risk of being exposed in case of large-scale data theft. The addition of DRM features would help organizations to effectively manage access to business-critical personal data held and controlled by Cloud storage. The risk of unencrypted content being exposed to unauthorized entities can be reduced by applying digital rights management technology (De Angelis & Di Marzo Serugendo, 2017; C. C. Lee et al., 2018; Munier et al., 2012; Torres et al., 2008).

Data integrity along with immutable licenses is also an important feature of DRM. This makes sure the content is not modified through ongoing communication over the Internet. As DRM encrypts the content data and applies access control mechanisms and policies to enforce legal access and to avoid unauthorized content consumption, it inherits the challenges of Cloud computing. Cloud computing suffers from new challenges of lightweight approach flexibility which makes it less suitable for Cloud DRM applications. This requires the Cloud DRM to meet the new requirement of multiple heterogeneous devices such as Smart TV, Smart Watches, Desktop, and many other distributed computing methods (Sicari et al., 2015). Keeping in view the increasing volume of audio, video, and other types of daily used media being stored in the Cloud, there is also a need to apply DRM as the second wall of defense for data security and in storage systems. Besides cloud-based solutions, some watermarking and proxy re-encryption-based solutions are proposed to protect the integrity of data in case of access delegation (HUANG et al., 2014; Nuñez et al., 2017; Verma & Singh, 2018). A standalone blockchain-based DRM with immutable properties is also proposed (Kishigami et al., 2015; Ma, Jiang, et al., 2018; Zhaofeng et al., 2018).

In all the solutions proposed, the strong encryption and integrity chain is complex to realize. However, efficient content protection is necessary to preserve the data ownership and copyright, thus preventing the content consumer entity from unauthorized consumption of the shared data in the Cloud. The integrity protection mechanism should also give a reasonable performance, low overheads, and lightweight operations to meet practical Cloud DRM requirements.

The data serving from Cloud technologies suffer from glitches and delays (Puliafito et al., 2019). Cloud DRM services that safeguard data, also need to provide optimal performance (Bedi et al., 2018; Chokngamwong & Jirabutr, 2015; Puliafito et al., 2019). This will serve the protected DRM content in a resource-constrained distributed computing environment. The latency is due to the geographically distributed nature of the spread of data in the Cloud. Fog computing is an extension of Cloud computing and is an emerging paradigm, proposed by the researcher, which covers this practical limitation of Cloud computing. Fog Computing improves the latency in Cloud computing technologies so that it satisfies the requirements of delay-sensitive applications. Therefore, it is important to discuss the Fog computing paradigm in Cloud DRM to serve protected data, such as geographically distributed healthcare information effectively (Zhang et al., 2015). Fog computing is hereby introduced to improve the service of delay-sensitive cloud applications, better mobility, and lower communication complexity by moving communication and storage near to the core network of the end-user. This would enable the devices and end-users to have better support for low latency application and mobility, by taking the communication to the edge of their networks.

In general, confidentiality, completeness, and availability of Internet platforms are still a problem with new and evolving online Internet technologies. Constructing scenarios of integrity protection with a focus on flexible protection as a social institution is an active research topic. Facebook users combine these behaviors to form privacy management strategies and intriguing differences among the privacy management strategies. These strategies differ in terms of perceptions and demographics like privacy concern, usefulness, enjoyment, and age, gender, and experience (Lankton et al., 2017). But DRM features offered by the current DRM technology landscape lack the flexibility to apply them which suits the wider application domain. This is because DRM is solving a complex problem. Even authorized users could tamper or fake the content and easily republish it. In a super re-distribution model such as social media, this can pose a threat to public reputation and propagate false information (Cheng et al., 2016). This problem is more prominent in online social media platforms and Cloud services due to the sharing feature and social networks are currently among the most popular services (Serrao et al., 2018). A DRM framework for online social networks (OSNs) was proposed (HUANG et al., 2014) to cater to the need for user privacy protections and control unauthorized content redistribution. Critical and confidential business documents are being stored and shared via online Cloud storage and social media platforms (Cook et al., 2017). The use of Cloud storage is very common today and therefore it is important to study DRM technology to enforce the copyrights policies over data and protect user data integrity in addition to the integrity protection feature provided by Cloud storage platforms. Controlled sharing is fundamental to distributed systems; yet, on the Web, and in the Cloud, sharing is still based on basic mechanisms (Joshi & Petrlic, 2013). Most DRM authorization mechanisms are not directly suited to the Cloud, since they are based on more expensive methods that can be difficult to implement and apply for long-lived, linkable identities and lightweight verifying of data integrity.

This chapter presents a thorough survey on the Digital Rights Management (DRM) landscape, its evolution, taxonomy, and state-of-the-art DRM techniques used by researchers. This article also adds the stakeholder's knowledge by highlighting the DRM deployments in the Fog computing service. However, Fog computing is tied to the IoTs, it is important to note that its use is applicable in several other contexts, e.g., content package, license management, key management, and content encryption and audit functions.

CLOUD-BASED DRM

Cloud-based DRM is the application of DRM technology to safeguard data stored and accessed from the Cloud. Cloud DRM manages the access control and authorization for its content and acts as a mechanism of enforcing the access policies. Cloud computing follows the N-services model in which the stored data could be consumed by diverse computational devices (e.g. Internet of Things devices, Desktop or Mobile, etc.) and thus provides easy and flexible access to the Cloud resources. Today is the era of Mobile technologies and with the use of high-speed internet, e.g. EDGE/3G/4G/5G, the mobile devices, though have limited memory, processor, and battery power are now trying to satisfy user performance expectations by offloading the processing to Cloud and thereby sharing their processing load to the Cloud. Consequently, there is an increase in depending on Cloud data storage providers to store the user's content. To effectively protect the contents by applying versatile DRM policies, many studies have been done by researchers to propose efficient DRM in the Cloud. We discussed those state-of-the-art techniques in the following section.

The recent increase in the utilization of Cloud computing has tremendously increased the use of data sharing features for various applications that lessened geographical distances by engaging tens of thousands of users, in real-time. Currently, both the private and Government sector enjoy massive data hosting in Cloud, and it's sharing as a mechanism to enhance organizational efficiency. Cloud-based DRM provides cloud consumers with content security services. Cloud-based DRM is much needed in today's IT revolution(Patranabis et al., 2017). Cloud computing has changed the paradigm of data services being used in academics, medical science, economics, e-commerce, and online social networking. In addition to the benefits of Cloud DRM solutions, many users sitting in different parts of the world could effectively collaborate, share, and exchange data with flexibility and convenience. The use of Cloud storage is currently increasing as individuals and organizations are relying on it for next-generation storage and collaboration purposes. In the generic Cloud DRM, the consumers are the end-user smart devices that utilize the protected content. DRM Cloud developer builds API for DRM Cloud services. The content owner provides the content, and the user has to request a usage license from the license server in Cloud. These entities constitute the application layer of Cloud DRM.

Despite all its benefits, the data stored in Cloud storage is vulnerable to data integrity, privacy, and security attacks, and also, it is not quite suitable for delay-sensitive applications like screaming and quick-response scenarios. The main strength of cloud-based DRM is its flexible and easy data sharing and access from multiple devices.

Cloud DRM inherits the challenges of both Cloud computing as well as efficient DRM. It is important to highlight the key characteristics of Cloud DRM to identify the research gap and issue of strength of Cloud DRM Following are the main characteristics of any Cloud DRM Security and Reliability.

CLOUD DRM AUTHORIZATION

Another study (Xie et al., 2021) attempts to address the content security or DRM in Cloud computing by focusing on attribute-based encryption (ABE). It uses attributes of users involved in encryption to apply DRM encryption. Also, it attempts to address the challenging requirement of multiple devices, for optimal handling in modern deployment scenarios. The author proposed hybrid cloud multi-authority ciphertext-policy attribute-based encryption (HCMACP-ABE). The proposed method utilizes Linear Secret-Sharing Schemes (LSSS) data structure for implementing a secure access method, which is independent of the private Cloud. The private Cloud is responsible for maintaining the overhead of the user's authorization list and verifying the user. Their approach targets the hybrid Cloud environment.

The aforementioned scheme empowers the content processing users and enforces the security of data by regulating the access for mobile devices in hybrid Cloud scenarios. Involved proxy layers and Cloud user assistance involve screening the request as a proxy layer as well as a component with application encryption and decryption. The aforementioned scheme applied Canetti's transformation as a reference for security and performance. The study aims to reduce the computation overheads, and also attempt to improve the efficiency of the mobile Cloud environment. However, an extra layer of Cloud user assistance (CUA) can bring additional cost and overhead in latency-sensitive scenarios. With the advent of Cloud computing in the public Cloud environment to facilitate the keywords of interest or search in protected content shared over public Cloud. A similar study (Cui et al., 2018) revolves around multi-user delegation and group search operation. They proposed an attribute-based keyword search with an efficient revocation scheme (AKSER). Their design improved the efficiency of user revocation and fine-grained authorization of the search and group authorized entities. Their method achieves semantic security, unlikability of keywords, and resilience to the collision.

In AKSER, content authors use personal or custom access policies to encrypt the file indices. That increase mapped over the role or category of the user authorized to query the index. The aforementioned approach focuses scalable multi-certificate authority access control mechanism capable of searching many keywords and many data owners. The resultant system relies upon the central server to implement efficient revocation operations per user basis. The aforementioned approach aims to improve the accuracy of cloud access control server in releasing the search use cases. Keeping the recent trend of using users attribute to the context of request as identity in the process of authorization, the authors (Lin, Wang, et al., 2021) presents identity-based encryption with an equality test (IBEET). The author proposed a new primitive, called identity-based encryption with equality test and timestamp-based authorization mechanism (IBEET-DBA). In the aforementioned approach, the content owner has the authority to control as well as validate data. The author addresses the limitation of ciphertext-specified authorization and user-specified authorization. The formal definition of the approach is primitive along with a security notion. Moreover, the author proposes the first IBEET-DBA scheme and demonstrates its security. Authors (Voundi Koe & Lin, 2019) redesigned proxy re-encryption to unlink users' personal data from the Cloud store thereby moving the identified mask from Cloud storage. Thereby leveraging the control of cloud storage for improved user privacy by enhanced authentication and authorization mechanism. This study also attempts to improve the flexibility of user authentication and authorization mechanism, and it saves the user from being online all the time to protect their data in the cloud. Another study by (Pareek & Purushothama, 2020) focuses on increasing the efficiency of Proxy Re-encryption (PRE) by requirements and also discusses its potential in solid versatile access control facilities. The study demonstrates that controlled sharing can be achieved efficiently with PRE for versatile delegation scenarios.

The authors (Deng et al., 2017) proposed Multi-user searchable encryption (MSE) which uses encryption and applies DRM to facilitate authorized users in searching for protected content. The proposed solution targets the use case involving Cloud storage where content is reshared for collaboration and the risk of unauthorized consumption but can leverage the search feature over the encrypted data. The study aims to address the practical limitations of Cloud authorization. Among the requirements highlighted are the security of data hiding and the ability to efficiently authorize or revoke a user to search over a file. The study highlights the gap that no existing scheme can achieve all these properties at the same time. The proposed schedule addresses the needs by applying attribute-based complex encryption operations which authorize other users to search for a subset of keywords in encrypted form. The proposed schedule uses an asymmetric bilinear map along with a keyword authentication organization binary tree (KABtree) to craft new ways to achieve performance benefits.

Another study (Antonolpoulos et al., 2018) focused on user privacy and automated Physical Access Control System (PACS). They propose to enhance the private Cloud capable of applying access control safely by encrypting sensitive information at the same time paying user privacy. The cloud service tracks the overall system activities in physical infrastructure and inbound alerts for a data breach or access violation. The approach involves processing logs in the public Cloud. The authors (Chadwick & Fatema, 2012) proposed cloud authorization using XACML applied to web services in Cloud. The author discusses the importance of simple policy implementation to handle the authorization and the design complexity. The authorization protocol needs to be application developer-friendly and as simple as possible, especially in the Infrastructure as a Service (IAAS) deployment. And the study proposes that the complexity needs to be under the web service interface, leaving less responsive infrastructure and simplicity. The proposed OASIS SAML-XACML is a solution to increase access control in the Cloud with less complexity toward the Cloud layer.

Parmar & Bhavsar, (2020) propose a new terminology RoT as an alternate terminology to address the unified need for Authentication, Access Control, Confidentiality, Scalability, Encryption, Integrity, and Authorization. The Author (Esposito, 2018) presents a model which aims to address the multistakeholder authorization among the organizations. The study focuses on the interoperable problem of authorization. The study also highlights the deficiency of effective support to enable the coexistence of multiple access control in a context. The study also advocates the need for dynamic approaches with greater support for seamless interaction of multiple roles with the cloud over time and to resist unauthorized data leak attacks. The aforementioned approach is based on ontology-based access control given the trust among entities of a process and also uses pseudonyms for privacy needs. The aforementioned approach proposes as a second-level defense for data going to the public cloud. That critical data is marked with the severity of access using tags that associate the trust. The research also highlights the access control beyond identity access scope by ensuring unlinkability of personal data or data which can facilitate the unauthorized consumption of data to predict user habits and profiling which is necessary for improved trust and transparency of digital systems.

Another study (Shen et al., 2017) addresses the computation overhead for data integrity in Cloud for resource-constrained environments. Authors present auditing mechanisms for Cloud storage auditing schemes suitable for cluster users and aiming to optimize the computation overhead from end-user devices. They named their proposed method as Third Party Medium (TPM). The TPM is in charge of generating authenticators for users and verifying data integrity on behalf of users. TPM is also specialized to enable end-users to make sure data is not modified by Cloud storage; this adds a significant contribution toward big problems of data integrity verification in the Cloud. TPM does time-consuming operations which users need to do thereby reducing the overhead and verifying data on end-user devices. This saves users from heavy decryption operations when interacting with its data saved in the Cloud. This results in user operations taken care of by the Cloud. The privacy is proposed to be extended by applying data blinding operations as users upload data. The authorization method is time-bound, making it easy for an authorized user to enjoy data integrity with time as additional parameters for Cloud data auditing needs. The security attached comprised content security areas such as content privacy infringing, data hacking, and unauthorized data access involved in the cloud layers. The study discovered that the attack gets more damaging as lower Cloud layers which directly relate to the OSI model and basic network service are directly affected (Hussain et al., 2017).

An effort focused on integrity protection features of Cloud DRM was made by the researcher (Lu et al., 2020) and their study emphasized the greater data integrity in shared data storage cloud servers. They attempt to address the integrity protection mechanism over the content by applying access control in mobile cloud computing. Their approach to integrity protection applies encryption over plain text data and aims to implement the second wall of protection to safeguard data privacy and integrity defense in case of data leaks. The approach attempts to optimize the secure and lightweight integrity verification scheme for Internet of Things (IoT) mobile terminal devices. They designed a data sharing method for data owners to share cloud data with authorized users. Finally, we proposed Merkle Hash Tree as a Version Based Merkle Hash Tree (VB_MHT) that can preserve the information of a block node fresh and improve security and integrity verification of the shared data. Their approach focuses on achieving lightweight operations of data owners. They also have defined mechanics for data collaboration, among authorized users, and sharing among users for downloading and consumption from shared cloud data. The author also presents the performance of computation and communication costs.

As the real-world deployment of IoT and Cloud will bring a high volume of data which will be hard to manage due to the scattered data spread over a distant geographic area. This challenge will give birth to mechanisms involving end-use devices and data owners in playing role in the infrastructure as a whole. The author (Tapas et al., 2020) presents IoT-Cloud based model for authorization and access delegation which also utilized blockchain technology. Although the study focuses on smart city requirements and presents smart contracts driven methods for smart features and assessed access need of control and delegation in IoT. Three real-world scenarios for access control and delegation in IoT use Blockchain technologies. They present a theoretical analysis of time and space complexity targeting create delegation, delete delegation, and check access. Their model was implemented on the Ethereum testnet

Ganache and public testnet Rinkeby. The study also presents the performance evaluation. Researchers (Bernal Bernabe et al., 2014) focused on the greater availability of the access control feature in cloud computing. The study focuses on increasing the adoption of the modern authorization model in practical cloud deployment scenarios. The author represents Role-based access control (RBAC), hierarchical objects (HO), conditional RBAC (cRBAC), and hierarchical RBAC (hRBAC) for cloud storage. The proposed model has support for multi-agent and federated access control features as well. The federated authorization together with semantic mapping of access is discussed. Resultant model address the fine-grained trust for administering a trusted foredated central server in cloud computing. The author has also presented the validation of the prototype by developing it using OpenStack with python and Java programming languages. The authors (Sun et al., 2020) presents searchable encryption scheme which makes up the defense of per user personalized linkable search. Their approach is using server-assisted searchable encryption. Multiple users are facilitated by selective authorization. The data owners only need to know the public key of an administration server to generate the searchable ciphertext. The study compares the related word in the parameters of search privacy, ease of use case, computational burden, and communication latency.

Another study (Sultan et al., 2018) highlights the future perspective of a secure inter-cloud authorization scheme called ICAuth. ICAuth uses ciphertext-policy attribute-based encryption (CP-ABE) for authorization of user access tokens. Their approach targets low latency, low communication overhead, and less storage consumption for lightweight computation costs. IAuth generates a single decryption key in a standalone manner independent of other entities. The one key can be used to access many resources. The revocation mechanism involved a re-encryption algorithm which has overhead in itself. ICAuth also aims to be more flexible and scalable for inter-cloud shared access scenarios. The author also presents security analysis and demonstrates that it is immune to Chosen Plaintext Attack (CPA). The performance analysis is presented keeping in view the use cases, network and file system overhead, latency, practical applicability, usability, and computation costs.

MAIN FOCUS OF THE CHAPTER

Issues, Controversies, Problems

In recent years, the research trend on Internet security, more focused on securing the Internet Backbone and Cloud as an infrastructure; however, the content security is a potential area to be researched further. The content security domain could help in securing the content in extreme data theft scenarios. For example, in the scenario of a data breach, after the data breach happened, the end-user data would be left as it is for malicious users. DRM could help to prevent unencrypted data disclosure by employing the content protection mechanism. Through this, DRM could also provide an extra layer of content security.

Current DRM landscapes have many content protection techniques. The popular one uses the methodologies of watermarking (Hou et al., 2018; Iftikhar et al., 2017; Kwon et al., 2016; Ma et al., 2016; Subramanyam et al., 2012), steganography (De Angelis & Di Marzo Serugendo, 2017; Mtech, 2015), image and video encryption (Thanh & Iwakiri, 2016) and Blockchain (Ma, Huang, et al., 2018; Ma, Jiang, et al., 2018). All of these methodologies have a bottleneck of performance and especially, in the case of large-scale deployment and serving such as Cloud computing. This is mainly because of the complex encoding and decoding algorithm. The small handheld devices have relatively low computational, storage, and internet capabilities. Therefore, the protected content distribution and consumption suffer from bad performance and less room for flexibility. This performance further degrades, when serving scattered and large amounts of time-sensitive critical data for Cloud. Currently, a large amount of users' data is being stored on Cloud based storage providers in a multi-service model. The strong encryption-based mechanism is insufficient to deliver the optimal DRM services (Singh et al., 2019). A recent study attempts to address this problem through efficient Key-Aggregate Cryptosystems with Broadcast DRM techniques for Cloud computing. Aggregate Keys Users would be able to decode many classes of data using only one key. That key is also of fixed size, so it could be easily sent to many users (Sachan et al., 2012). Many researchers have proposed DRM technologies, so that they would be suitable for heterogeneous smart devices and the interoperability among various DRM technologies (H. Lee et al., 2016a). These techniques lack practicality, and the proposed solutions are not applicable to meet the next generation performance needed. Therefore, more stagnant Cloud DRM technologies are needed. Currently, the Fog paradigm is rapidly adopted to provide multi-tier, on-demand, flexible, and cost-effective services to users. Legacy client-server data and database hosting are being replaced with Cloud hosting, to enjoy these benefits. Fog computing will solve latency problems in Cloud computing which will in turn phase out lightweight cloud DRM. Hence DRM in the Cloud would grow more and more important; especially for protecting content (Hu et al., 2017). The trustworthiness of the computing unit which is responsible to process the data is also important and could be assessed if data integrity is safe. As in, software as a service (SaaS) model, the computing devices of service providers consume the critical and confidential data of the source organization. The data source organization loses direct control over the data they provide (Zafar et al., 2017). This leaves data owners at risk of data theft or data misuse. Data source organization is not responsible for third-party managing and using their data, it also improves legal and data confidentiality concerns. In service-oriented architecture, it is possible to authenticate the integrity of untrusted middleware data processing elements (J. Huang et al., 2014). Due to this data immutability demand has gained as security by design. The research on Blockchain-based methods for leveraging content immutable benefits is growing. The proposed methods utilize elliptic curve encryption and heavy miner networks to safeguard data integrity and safeguard data tempering when access is delegated. However, the Blockchain-based solution for DRM lacks partiality and faces performance challenges. Implementation of distributed systems is difficult and Cloud computing is a stack of technologies so more efforts are needed to design lightweight Cloud DRM methods.

Message authentication code (MAC), sometimes known as a tag, is a method used to authenticate a message. The specific type of message authentication code (MAC) involving a cryptographic hash function and a secret cryptographic key is Hashed Message authentication code (HMAC). Macaroon (Birgisson et al., 2014) are HMAC based bearer tokens that were first presented in 2014. Their mechanism allows them to create an immutable signature to help delegate Internet resources just like cookies used in internet browsers. However, Macaroon is more flexible and lightweight at the same time it's immutable which means its tamper-proof in nature as it builds chain of authorization proof of data and signatures. This chain of lightweight signatures has been found suitable to implement lightweight and robust DRM for large distributed systems. Macaroon's features of tamper-proof and ease of integration makes it a good candidate for Cloud DRM in secure content sharing and controlled content management in collaboration scenarios.

Even More Issues, Controversies, Problems

In the past decade, Cloud computing adoption has been significant as remote, on-demand, cost-effective and revolutionary data storage. This adoption has overwhelmingly increased the volume of personal and business-critical data stored in these cloud storages which is often shared for collaboration. This data sharing required additional access control technologies supporting the cloud benefits. However, cloud DRM technologies lack flexible and lightweight solutions. This is because the traditional authorization mechanisms lack practical aspects. The multi-layer Cloud DRM architecture with multi-party interaction couldn't be satisfied by the static token-based authorization mechanism. The approach in literature Uses methods that add an extra level of complexity and performance tradeoff, in realizing lightweight authorization solutions with immutable policy.

This research aims to address the need for flexible and tamper-proof licenses, by proposing an alternative approach for Cloud DRM. The proposed approach provides strong license integrity, flexible re-attenuation of license, and lightweight implementation for protecting and verifying data integrity in Cloud DRM.

SOLUTIONS AND RECOMMENDATIONS

This study represents Macaroons as an alternative method for lightweight, flexible, and tamper-proof integrity in releasing immutable integrity protection methods and flexible policies for DRM in the Cloud. The proposed method is analyzed and critically discusses how the Macaroon approach could achieve tamper-proof and flexible authorization in the Cloud for an efficient DRM solution. Macaroon-based solutions are constructed using HMAC chain. The Macaroon feature with the flexibility to add unlimited caveats and its temper-proof nature of HMAC chain generate lightweight credentials that are better than the existing authorization techniques used in Cloud DRM for flexible and efficient spatial data management and related constraints. This study also analyzes the operation cost of Macaroon by benchmarking its libraries. The benchmark result correlates the implementation technology with the performance of sub-operations.

Significant research is being carried out to improve the security of transport layer protocol, such as the Public Key Infrastructure (PKI). This thesis aims to improve content security at the application layer by proposing Macaroon as a method for Cloud DRM solutions. We achieved the objective outlined in section 1.4 by reviewing the problems of Cloud DRM and identified the state-of-the-art constraints and challenges in designing DRM technologies better and a flexible way of protecting the intellectual property of online content in a distributed computing environment.

IMPLICATION OF RESEARCH

Macaroon has a chain of tamper-proof authorization with lightweight verification and is found to be suitable for flexible authorization of content in Cloud DRM. The integrity of content in Cloud DRM is an important problem to address with the flexible and lightweight operation. Macaroon as a method for the integrity chain mechanism in the Cloud would be a lightweight solution to the problem that other studies are trying to achieve: Macaroon as Authorization Bearer Token Controlled with more flexible rights association. Therefore, there is a need for bearer tokens that tamper with proof to some considerable degree. In this study, we propose a tamper-proof integrity protection solution for Cloud DRM leveraging Macaroon. The tamper-proof nature of Macaroon caveats makes it an important feature of secure DRM systems which resist against various attacks. Hence, focusing on the solution perspective, we have designed and implemented all the protocols involving a digital token from its generation and storage to circulation between different players, and final redemption.

We have analyzed the benefits of being efficient, flexible, and immutable from a DRM perspective. The sub-operation of the Macaroon creation, verification, etc. gives benefits of immutable proof of access with much flexibility to further limit the access in a standalone manner. The flexibility also brings lightweight. The evaluation of Macaroon approach for Cloud DRM and estimate the cost of Macaroon sub-operations and comparison with the implementation of different programming language Macaroons applicability to large and open distributed systems for content protection and performance comparison of various Macaroon implementation and Macaroon approach is found flexible.

ANALYSIS OF SALIENT WORK

The study started with a systematic literature review of the DRM technologies being proposed in the DRM. The generic DRM system we have taken as a reference aims to support flexible applications with Cloud DRM features. The systematic literature review to understand the useful methods and protocols to solve the said problem. The literature was mainly gathered from applying search queries in an academic journal and conference databases with the relevant keywords, such as digital rights management, Cloud DRM, Cloud authentication, and decentralized authorization. The downloaded literature was studied in depth in summarizing the approach and future challenges.

In the beginning, we established some standards for a new DRM Cloud system for better performance and security. This study also discusses Fog Computing in DRM to solve the latency problems in Cloud DRM.

RESEARCH GAP AND ANALYSIS

As millions of users use social media to share pictures, videos, and other daily life documents. This Sharing is a very common and useful way of sharing personal contacts and information. What makes these social services so unique and attractive to users is not the fact that they allow them to know other persons, but the fact that they allow the users to expose their network of friends to others and share their content and ideas. But the user has very limited control over its data after the data is available on social networking platforms whereas the platform and authorizer user can tamper with the data and easily republish it.

The success of online social networking platforms is that they allow users to share their social network with others and also can see the social network of others. The targeted audience to see the shared content can be the direct connection or the further links of our direct connected users or anyone on the same platform with a valid account (regardless of whether it's your connection or not). These social network sharing functionalities are very important and inviting of further social interaction however, they are at the same time, the cause of serious privacy and security concerns. Currently, the sharing control is not on the end-user side, but on the social platform side with a limited set of features for users to protect their privacy and intellectual property. Ideally, users should be able to handle their content and define the boundaries and rules of access to it. This is something that does not yet happen in most current social networks.

Most online social networking platforms use the Cloud computing paradigm to store and serve data. After a user transfers the data to any Cloud-based storage provider, the data itself is managed by the Cloud provider and the data owner has little control over its data in rest or when the user shares the data with another user or with the online social media platforms. DRM in the Cloud refers to the technology which enables content publishers to access control of their content for the content-consuming devices. DRM in the Cloud is necessary, to preserve intellectual property, to manage data confidentiality, data integrity, and data security, especially after content is shared with the third party. This third-party verification feature is an emerging topic of research and optimal access technology can help to solve the problem in Cloud DRM.

In the literature that Attribute-based Encryption, Advance Watermarking (Ma et al., 2016), DRM relying on Trusted third party (TTP) (Birrell & Schneider, 2013), DRM Without relying on TTP (Win et al., 2012), and Blockchain (Public verified)(Ma, Jiang, et al., 2018) based DRM were proposed, however the Message authentication code approach to DRM was not much explored. As the Macaroon approach can provide easy and practical authorization, therefore this study analyzed the Macaroon approach to DRM for a flexible and tamper-proof solution for data integrity in distributed computing.

Macaroons DRM could provide a better solution as it provides better-delegated authorization, autonomous attenuation, and scalable cross-domain. Storage providers are now exploring how to use macaroons (Birgisson et al., 2014). Other storage systems explore macaroons (SurfSARA, MinE, dCache, and SWESTORE) (Millar et al., 2018), Fast revocation, and Carry its cryptographic signature as proof for access.

EVALUATION OF THE MACAROON'S OPERATION USED IN DRM

Macaroon carries its proof of authorization as a chain of delegations. In this way, the delegated authority enjoys the same level of access and without involving a central authority. Accessing a Macaroon grants you any permissions, appending more caveats (conditions) to it to further squeeze the permissions one wants to authorize, and then send the reduced Macaroon to the targeted receiving party. To use the delegated permissions, the attenuated Macaroon can be presented to the service provider. The service provider can verify whether the nonce of the caveats in the Macaroon is modified or not as originally issued by the service provider. The benchmarking results aim to evaluate the Macaroon operations cost (time) in different programming languages to quantify the optimal deployment scenario of Macaroon in Cloud DRM. The Macaroon system model and benchmarking results will be useful in selecting a technology stack for DRM implementations and set directions for further enhancement of the Macaroon library, its standardization and increased adaptability for enhanced use.

CHALLENGES OF MACAROON APPROACH

Macaroons are developed based on the assumption that the service wishes to give access permissions to another service and both are already communicating with each other. Macaroon also doesn't care about

Table 1. Macaroon Cloud DRM comparison with traditional approaches

DRM Feature	Related Studies	Macaroon Based DRM in Cloud Computing
Multi-user searchable encryption (MSE)	(Deng et al., 2017)	The multi-user search is not supported
Focus of ABE	(Xie et al., 2021), (Cui et al., 2018; Lin, Wang, et al., 2021; Pareek & Purushothama, 2020; Sultan et al., 2018; Voundi Koe & Lin, 2019)	Unlimited distinct contextual attributes could be added and supported
hybrid cloud environment	(Esposito, 2018; Hussain et al., 2017; Sultan et al., 2018; Tapas et al., 2020; Xie et al., 2021)	Better in terms of complexity
aims to reduce the computation overheads	(Cui et al., 2018; Deng et al., 2017; Lu et al., 2020; Shen et al., 2017; Sultan et al., 2018; Tapas et al., 2020; Xie et al., 2021)	Better in terms of Lightweight and chain of HMAC make it more suitable for resource constrained environments.
improve the efficiency	(Bernal Bernabe et al., 2014; Esposito, 2018; Lin, Sun, et al., 2021; Lin, Wang, et al., 2021; Lu et al., 2020; Shen et al., 2017; Sultan et al., 2018; Tapas et al., 2020; Xie et al., 2021)	The method is more flexible and brings more performance and implementation efficiency
Physical Access Control System (PACS).	(Antonolpoulos et al., 2018)	Not supported
Simple and flexible approach	(Chadwick & Fatema, 2012; Cui et al., 2018; Voundi Koe & Lin, 2019)	Unlimited attenuation makes it a far more flexible method
Integrity protection	(Lu et al., 2020; Shen et al., 2017)	Bring immutability by design into cloud DRM

Table 2. Comparison of synchronous and asynchronous e-learning in growing digital ethics demand era

	Synchronous E-Learning	Asynchronous E-Learning
When	● Discussing less complex issues. ● Getting acquainted. ● Planning tasks.	● Reflecting on complex issues. ● When synchronous classes cannot be attended due to illness, work, family or other commitments.
Why	● Students become more committed and motivated due to getting quick responses.	● Students have more time to reflect as the quick response is not immediately expected.
How	● In addition to face-to-face class, various synchronous means including video conferencing, instant messaging and conversation (chat) are used.	● Various asynchronous means such as e-mail, discussion boards, and blogs are used.
Online	Synchronous means: ● Virtual Classroom. ● Video/teleconferencing. ● Conversation (chat) rooms/instant messaging.	Asynchronous means: ● Web-based teaching/ computer based teaching. ● Threaded discussion groups. ● Recorded live events. ● Online documents/ email/global announcement.
Offline	Synchronous means: ● Face to face classroom. ● Hands on laboratory practices. ● Field trips, field work.	Asynchronous means: ● Bound books/ learning resources. ● Videos/Echo360/Lectopia. ● Audio tapes.
Examples	● Students work in groups and can use instant messaging as a support for getting to know each other, exchanging ideas, and planning tasks. ● A teacher who wants to present concepts from the literature in a simplified way might give an online lecture by video conferencing.	● Students expected to reflect individually on course topics may be asked to maintain a blog. ● Students are expected to share reflections regarding course topics and critically assess their peers' ideas. They may be asked to participate in online discussion on a discussion board.

the underlying channel and assumes it to be secure. It is only dealing with the delegation of the existing permissions of a service. These features of Macaroon make it work as a standalone and flexible discovery mechanism. The verification method of macaroons also sets up an engineering challenge in managing secrets. As the minting and reconstruction of a Macaroon requires the root secret key, this inevitably means one of two things. Either the minting or verifying is done solely by that one service or the secret needs to be shared, which in turn requires methods for secure distribution and management of secrets. Macaroons are designed for short-time permissions and the order of access permission doesn't matter much. Macaroons heavily depend on service providers using it. Macaroon works in the same context as the service Cloud provider gave at the beginning of granting access and the level of access is attenuated as its delegate. Every time a Macaroon is attenuated, it transforms into another Macaroon. Only the initial service provider can verify these delegated accesses to Macaroon, similar to that of Simple public key infrastructure (SPKI). In both of these systems, only the service provider can verify permissions. Macaroons constitute a simpler system and have performance benefits. We summarize the discovered pros and cons of Macaroon in table 3.

Table 3. Advantage and disadvantages of macaroon as authorization token

Advantages	Challenges
Allows only a secure HMAC option	Formalization of the logic needed
Third-party caveats enable novel options for authentication and authorization	Lack of interoperability and platform dependency
Security reliant on underlying one-way hash function used as a black box	Minting and verification using symmetric cryptography creates secrets management challenges
Attenuation and delegation of access allows for more flexibility	Lack of standard implementation leaves a lot of responsibility for the developer
Enables granular resource-level access control based on the set authorization policy	The more you want to do with them, the more logic needs to be implemented on the application level

FUTURE RESEARCH DIRECTIONS

Digital content is always at risk of being manipulated and misused. Finding new and robust and tamper-proof authentication and authorization methods are important for a safer and secure Internet. This work experimented on several real-world Cloud content abuse scenarios; especially the social media content which can be used for data theft, intellectual property rights violation, etc.

The development of an advanced version of the DRM System suitable for Cloud computing would enable researchers to protect online content more effectively, with greater content security. The integrity of online content lacks some features compared to the real adoption and deployment challenges, such as interoperability, internationalization, Internet policy, etc. Research on this issue would benefit the DRM research community in the future. (Xie et al., 2021).

CONCLUSION

Content security in Cloud storage and easy unauthorized further dissemination potentially help in ease out diversity and involution while it also brings emerging challenges for Cybersecurity, for the healthy accelerating the adoption and a safer Internet space. The current industry 4.0 revolution, as well as adoption and acceleration efforts are leveraging Cloud computing as a means to store their Cyber-physical data. This makes Industry 4.0 infrastructure, as well as Cloud content susceptible to content abuse, which may lead to data theft. The damage due to their unauthorized data leaks is increasing. Digital Rights Management (DRM) practices for better and ethically safe online space are becoming more and more important. This chapter highlights some of the key issues of distributed authorization in DRM from a technical perspective of performance, flexibility, and immutability features. This chapter will help to streamline the Internet governance implementation roadmap for the industry 4.0 revolution. Future studies may like to address the data confidentiality and enhanced ethical usage of Cloud in DRM.

REFERENCES

Antonolpoulos, F., Petrakis, E. G. M., Sotiriadis, S., & Bessis, N. (2018). A physical access control system on the cloud. *Procedia Computer Science*, *130*, 318–325. doi:10.1016/j.procs.2018.04.045

Bedi, R. K., Singh, J., & Gupta, S. K. (2018). MWC: An efficient and secure multi-cloud storage approach to leverage augmentation of multi-cloud storage services on mobile devices using fog computing. *The Journal of Supercomputing*. Advance online publication. doi:10.100711227-018-2304-y

Bernal Bernabe, J., Marin Perez, J. M., Alcaraz Calero, J. M., Garcia Clemente, F. J., Martinez Perez, G., & Gomez Skarmeta, A. F. (2014). Semantic-aware multi-tenancy authorization system for cloud architectures. *Future Generation Computer Systems*, *32*, 154–167. doi:10.1016/j.future.2012.05.011

Birgisson, A., Gibbs Politz, J., Erlingsson, Ú., Taly, A., Vrable, M., & Lentczner, M. (2014). *Macaroons: Cookies with Contextual Caveats for Decentralized Authorization in the Cloud*. doi:10.14722/ndss.2014.23212

Birrell, E., & Schneider, F. B. (2013). Federated Identity Management Systems: A Privacy-Based Characterization. *IEEE Security and Privacy*, *11*(5), 36–48. doi:10.1109/MSP.2013.114

Chadwick, D. W., & Fatema, K. (2012). A privacy preserving authorisation system for the cloud. *Journal of Computer and System Sciences*, *78*(5), 1359–1373. doi:10.1016/j.jcss.2011.12.019

Cheng, Y. Y., Li, H., & Zhang, N. (2016). Character-Based Online Key Management in Cloud Computing Environment. *Proceedings of 2016 Ieee Advanced Information Management, Communicates, Electronic and Automation Control Conference (Imcec 2016)*, 738–741. 10.1109/IMCEC.2016.7867307

Chokngamwong, R., & Jirabutr, N. (2015). Mobile Digital Right Management with Enhanced Security using Limited-Use Session Keys. *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Ecti-Con)*.

Cook, A., Robinson, M., Ferrag, M. A., Maglaras, L., He, Y., Jones, K., & Janicke, H. (2017). *Internet of Cloud: Security and Privacy issues*. Academic Press.

Cui, J., Zhou, H., Zhong, H., & Xu, Y. (2018). AKSER: Attribute-based keyword search with efficient revocation in cloud computing. *Information Sciences*, *423*, 343–352. doi:10.1016/j.ins.2017.09.029

Das, A. K., Mishra, D., & Mukhopadhyay, S. (2015). An anonymous and secure biometric-based enterprise digital rights management system for mobile environment. *Security and Communication Networks*, *8*(18), 3383–3404. doi:10.1002ec.1266

De Angelis, F. L., & Di Marzo Serugendo, G. (2017). SmartContent—Self-Protected Context-Aware Active Documents for Mobile Environments. *Electronics (Basel)*, *6*(1), 17. Advance online publication. doi:10.3390/electronics6010017

Deng, Z., Li, K., Li, K., & Zhou, J. (2017). A multi-user searchable encryption scheme with keyword authorization in a cloud storage. *Future Generation Computer Systems*, *72*, 208–218. doi:10.1016/j.future.2016.05.017

Esposito, C. (2018). Interoperable, dynamic and privacy-preserving access control for cloud data storage when integrating heterogeneous organizations. *Journal of Network and Computer Applications*, *108*, 124–136. doi:10.1016/j.jnca.2018.01.017

Hou, J. U., Kim, D., Ahn, W. H., & Lee, H. K. (2018). Copyright Protections of Digital Content in the Age of 3D Printer: Emerging Issues and Survey. *IEEE Access: Practical Innovations, Open Solutions*, *6*, 44082–44093. doi:10.1109/ACCESS.2018.2864331

Hu, P., Dhelim, S., Ning, H., & Qiu, T. (2017). Survey on fog computing: architecture, key technologies, applications and open issues. Journal of Network and Computer Applications. doi:10.1016/j.jnca.2017.09.002

Huang, J., Lu, P., Juang, W., Fan, C., Lin, Z., & Lin, C. (2014). Secure and efficient digital rights management mechanisms with privacy protection. *Journal of Shanghai Jiaotong University (Science)*, *19*(4), 443–447. doi:10.100712204-014-1523-5

Huang, Q., Fu, J., Ma, Z., Yang, Y., & Niu, X. (2014). Encrypted data sharing with multi-owner based on digital rights management in online social networks. *Journal of China Universities of Posts and Telecommunications*, *21*(1), 86–93. doi:10.1016/S1005-8885(14)60273-9

Huang, Q., Ma, Z., Fu, J., Niu, X., & Yang, Y. (2013). *Attribute Based DRM Scheme with Efficient Revocation in Cloud Computing* (Vol. 8). doi:10.4304/jcp.8.11.2776-2781

Hussain, S. A., Fatima, M., Saeed, A., Raza, I., & Shahzad, R. K. (2017). Multilevel classification of security concerns in cloud computing. *Applied Computing and Informatics, 13*(1), 57–65.

Iftikhar, S., Kamran, M., Munir, E. U., & Khan, S. U. (2017). A Reversible Watermarking Technique for Social Network Data Sets for Enabling Data Trust in Cyber, Physical, and Social Computing. *IEEE Systems Journal*, *11*(1), 197–206. doi:10.1109/JSYST.2015.2416131

Joshi, N., & Petrlic, R. (2013). Towards practical privacy-preserving digital rights management for cloud computing. *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC),* 265–270. 10.1109/CCNC.2013.6488456

Kishigami, J., Fujimura, S., Watanabe, H., Nakadaira, A., & Akutsu, A. (2015). The Blockchain-based Digital Content Distribution System. *Proceedings 2015 Ieee Fifth International Conference on Big Data and Cloud Computing Bdcloud 2015*, 187–190. 10.1109/BDCloud.2015.60

Koulouzis, S., Mousa, R., Karakannas, A., de Laat, C., & Zhao, Z. (2018). Information Centric Networking for Sharing and Accessing Digital Objects with Persistent Identifiers on Data Infrastructures. *Proceedings of the 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 661–668. 10.1109/CCGRID.2018.00098

Kwon, G. R., Lama, R. K., Pyun, J. Y., & Park, C. S. (2016). Multimedia digital rights management based on selective encryption for flexible business model. *Multimedia Tools and Applications*, *75*(12), 6697–6715. doi:10.100711042-015-2563-z

Lankton, N. K., McKnight, D. H., & Tripp, J. F. (2017). Facebook privacy management strategies: A cluster analysis of user privacy behaviors. *Computers in Human Behavior*, *76*, 149–163. doi:10.1016/j.chb.2017.07.015

Lee, C. C., Li, C. T., Chen, Z. W., Lai, Y. M., & Shieh, J. C. (2018). An improved E-DRM scheme for mobile environments. *Journal of Information Security and Applications*, *39*, 19–30. doi:10.1016/j.jisa.2018.02.001

Lee, H., Park, S., Seo, C., & Shin, S. U. (2016a). DRM cloud framework to support heterogeneous digital rights management systems. *Multimedia Tools and Applications*, *75*(22), 14089–14109. doi:10.100711042-015-2662-x

Lin, X.-J., Sun, L., Qu, H., & Zhang, X. (2021). Public key encryption supporting equality test and flexible authorization without bilinear pairings. *Computer Communications*, *170*, 190–199. doi:10.1016/j.comcom.2021.02.006

Lin, X.-J., Wang, Q., Sun, L., & Qu, H. (2021). Identity-based encryption with equality test and datestamp-based authorization mechanism. *Theoretical Computer Science*, *861*, 117–132. doi:10.1016/j.tcs.2021.02.015

Lu, X., Pan, Z., & Xian, H. (2020). An integrity verification scheme of cloud storage for internet-of-things mobile terminal devices. *Computers & Security*, *92*, 101686. doi:10.1016/j.cose.2019.101686

Ma, Z. F., Huang, J. Q., Jiang, M., & Niu, X. X. (2016). A Novel Image Digital Rights Management Scheme with High-Level Security, Usage Control and Traceability. *Chinese Journal of Electronics*, *25*(3), 481–494. doi:10.1049/cje.2016.05.014

Ma, Z. F., Huang, W. H., Bi, W., Gao, H. M., & Wang, Z. (2018). A Master-Slave Blockchain Paradigm and Application in Digital Rights Management. *China Communications*, *15*(8), 174–188. doi:10.1109/CC.2018.8438282

Ma, Z. F., Jiang, M., Gao, H. M., & Wang, Z. (2018). Blockchain for digital rights management. *Future Generation Computer Systems-the International Journal of Escience*, *89*, 746–764. doi:10.1016/j.future.2018.07.029

Millar, A. P., Adeyemi, O., Behrmann, G., Fuhrmann, P., Garonne, V., Litvinsev, D., Mkrtchyan, T., Rossi, A., Sahakyan, M., & Starek, J. (2018). Storage for Advanced Scientific Use-Cases and Beyond. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP),* 651–657. 10.1109/PDP2018.2018.00109

Mtech, R. K. (2015). *The Non-Tangible Masking of Confidential Information using Video Steganography*. Academic Press.

Munier, M., Lalanne, V., & Ricarde, M. (2012). Self-Protecting Documents for Cloud Storage Security. *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, 1231–1238. 10.1109/TrustCom.2012.261

Nuñez, D., Agudo, I., & Lopez, J. (2017). Proxy Re-Encryption: Analysis of constructions and its application to secure access delegation. *Journal of Network and Computer Applications*, *87*, 193–209. doi:10.1016/j.jnca.2017.03.005

Pareek, G., & Purushothama, B. R. (2020). Proxy re-encryption for fine-grained access control: Its applicability, security under stronger notions and performance. *Journal of Information Security and Applications*, *54*, 102543. doi:10.1016/j.jisa.2020.102543

Parmar, P., & Bhavsar, M. (2020). Achieving Trust using RoT in IaaS Cloud. *Procedia Computer Science*, *167*, 487–495. doi:10.1016/j.procs.2020.03.264

Patranabis, S., Shrivastava, Y., & Mukhopadhyay, D. (2017). Provably Secure Key-Aggregate Cryptosystems with Broadcast Aggregate Keys for Online Data Sharing on the Cloud. *IEEE Transactions on Computers*, *66*(5), 891–904. doi:10.1109/TC.2016.2629510

Puliafito, C., Mingozzi, E., Longo, F., Puliafito, A., & Rana, O. (2019). Fog Computing for the Internet of Things: A Survey. *ACM Trans. Internet Technol., 19*(2), 18:1-18:41. doi:10.1145/3301443

Sachan, A., Emmanuel, S., & Kankanhalli, M. S. (2012). Aggregate Licenses Validation for Digital Rights Violation Detection. ACM Trans. *Multimedia Comput. Commun. Appl., 8*(2S), 37:1-37:21. doi:10.1145/2344436.2344443

Salman, O., Elhajj, I., Chehab, A., & Kayssi, A. (2018). IoT survey: An SDN and fog computing perspective. *Computer Networks*, *143*, 221–246. doi:10.1016/j.comnet.2018.07.020

Serrao, C., Marques, J., Dias, M., & Delgado, J. (2018). *Open-source software as a driver for digital content e-commerce and DRM interoperability*. Academic Press.

Shen, W., Yu, J., Xia, H., Zhang, H., Lu, X., & Hao, R. (2017). Light-weight and privacy-preserving secure cloud auditing scheme for group users via the third party medium. *Journal of Network and Computer Applications*, *82*, 56–64. doi:10.1016/j.jnca.2017.01.015

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, *76*, 146–164. doi:10.1016/j.comnet.2014.11.008

Singh, A. K., Nag, A., Karforma, S., & Mukhopadhyay, S. (2019). Implementation of multi-agent based Digital Rights Management System for Distance Education (DRMSDE) using JADE. *International Journal of Advanced Computer Science and Applications*, *10*(3), 343–352. doi:10.14569/IJACSA.2019.0100345

Subramanyam, A. V., Emmanuel, S., & Kankanhalli, M. S. (2012). Robust Watermarking of Compressed and Encrypted JPEG2000 Images. *IEEE Transactions on Multimedia*, *14*(3), 703–716. doi:10.1109/TMM.2011.2181342

Sultan, N. H., Barbhuiya, F. A., & Laurent, M. (2018). ICAuth: A secure and scalable owner delegated inter-cloud authorization. *Future Generation Computer Systems*, *88*, 319–332. doi:10.1016/j.future.2018.05.066

Sun, L., Xu, C., Li, C., & Li, Y. (2020). Server-aided searchable encryption in multi-user setting. *Computer Communications*, *164*, 25–30. doi:10.1016/j.comcom.2020.09.018

Tapas, N., Longo, F., Merlino, G., & Puliafito, A. (2020). Experimenting with smart contracts for access control and delegation in IoT. *Future Generation Computer Systems*, *111*, 324–338. doi:10.1016/j.future.2020.04.020

Thanh, T. M., & Iwakiri, M. (2016). Fragile watermarking with permutation code for content-leakage in digital rights management system. *Multimedia Systems*, *22*(5), 603–615. doi:10.100700530-015-0472-7

Torres, V., Serrao, C., Dias, M. S., & Delgado, J. (2008). Open DRM and the future of media. *IEEE MultiMedia*, *15*(2), 28–36. doi:10.1109/MMUL.2008.38

Verma, G. K., & Singh, B. B. (2018). Efficient identity-based blind message recovery signature scheme from pairings. *IET Information Security*, *12*(2), 150–156. doi:10.1049/iet-ifs.2017.0342

Voundi Koe, A. S., & Lin, Y. (2019). Offline privacy preserving proxy re-encryption in mobile cloud computing. *Pervasive and Mobile Computing*, *59*, 101081. doi:10.1016/j.pmcj.2019.101081

Win, L. L., Thomas, T., & Emmanuel, S. (2012). Privacy Enabled Digital Rights Management Without Trusted Third Party Assumption. *IEEE Trans. Multimedia, 14*(3–1), 546–554. doi:10.1109/TMM.2012.2189983

Xie, M., Ruan, Y., Hong, H., & Shao, J. (2021). A CP-ABE scheme based on multi-authority in hybrid clouds for mobile devices. *Future Generation Computer Systems*, *121*, 114–122. doi:10.1016/j.future.2021.03.021

Xu, R. Z., Zhang, L., Zhao, H. W., & Peng, Y. (2017). Design of Network Media's Digital Rights Management Scheme Based on Blockchain Technology. *2017 IEEE 13th International Symposium on Autonomous Decentralized Systems (Isads 2017),* 128–133. 10.1109/ISADS.2017.21

Zafar, F., Khan, A., Suhail, S., Ahmed, I., Hameed, K., Khan, H. M., Jabeen, F., & Anjum, A. (2017). Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes. *Journal of Network and Computer Applications*, *94*, 50–68. doi:10.1016/j.jnca.2017.06.003

Zhang, Z. Y., Wang, Z., & Niu, D. M. (2015). A novel approach to rights sharing-enabling digital rights management for mobile multimedia. *Multimedia Tools and Applications*, *74*(16), 6255–6271. doi:10.100711042-014-2135-7

Zhaofeng, M., Weihua, H., & Hongmin, G. (2018). A new blockchain-based trusted DRM scheme for built-in content protection. *Eurasip Journal on Image and Video Processing, 2018*(1). doi:10.1186/s13640-018-0327-1

Chapter 4
Keeping the UN Convention on the Rights of the Child Relevant in the Digital Age

Susan E. Zinner
Indiana University Northwest, USA

ABSTRACT

This chapter considers how the U.N. Convention on the Rights of the Child, adopted by the U.N. General Assembly in 1989 and ratified by every nation except the United States, protects the present and future rights of all children. However, the digital rights of children could not have been anticipated when the treaty was drafted. How should parents, legislators, child advocates, and others strive to both protect children from potential internet harm while still allowing children to develop the requisite skills needed to negotiate the internet alone? How best to achieve the balance between protection and digital participation will be the primary focus of this chapter.

INTRODUCTION

The UNCRC and the Rights of Children

The internet offers opportunities for children around the world to connect with each other and to learn from each other. While parents may fear the risk that these interactions pose to children, some authors have suggested that harsh policies designed to limit internet access by children are created, at least in part, to reduce adult anxiety (Vickery, 2017). In the United States, there was a great deal of concern about the risks posed by “stranger danger” that children might meet on line (and possibly, later, in person). However, the overwhelming risk by individuals to children come from those at home and not strangers (Vickery, 2017).

While parents and society have traditionally opted to protect children from risks both inside and outside the home, protection and acknowledgement of rights did not invariably go hand-in-hand. Until very recently few people felt that children deserved any rights at all (Wall, 2017). Adults, who may be

DOI: 10.4018/978-1-7998-8467-5.ch004

only too willing to recognize the cognitive limitations imposed by childhood on some children, may be unwilling to acknowledge the cognitive maturity of others. Since children mature emotionally and cognitively at different rates—just as their physical developmental rates vary—many societies have traditionally opted to err on the side of caution and limit their exposure to individuals, substances and media deemed risky. They are, in effect, protecting them from their immature decision-making abilities. However, this approach may have the unintended side effect of leaving that child ill-prepared for adulthood and independent decision-making.

The U.N. Convention on the Rights of the Child (UNCRC) suggests that children possess the requisite cognitive maturity to either make their own decisions or to participate in the decision-making process. In the inherent tension between protection of vulnerable children and allowing children to fully participate in decisions, including decisions in the digital world, the spirit guiding the UNCRC advocates for participation. The central role that the digital world now plays in the lives of all global citizens today, including children could not have been imagined.

This chapter will explore how the UNCRC would likely address the right to digital access by children, how adults in the lives of children should balance the protection and participation rights of children seeking to exercise these rights, what guidance the “evolving capacities” standard in the UNCRC provides in the context of minors exercising participation rights, how to respond to concerns about unequal digital media access, issues of good child citizenship involving digital media and international laws and reports that may provide some guidance in resolving these issues.

THE EVOLUTION OF THE UNCRC

The UNCRC evolved from several historical documents dating back to World War I. From the 1924 Geneva Declaration of the Rights of the Child to the 1959 Declaration on the Rights of the Child, the international community has advocated for children for a century. The Preamble to the UNCRC notes that “recognition of the *inherent dignity* and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world” (United Nations Office of the High Commissioner [OHCHR], 1989). The UNCRC begins with the inherent dignity present in all individuals and suggests that equity in all aspects of the lives of children is a goal of this document; inequities in power structures such as the law, health systems, political structures and other factors amenable to change are considered. Eventually the UNCRC was unanimously adopted by the U.N. General Assembly in 1989, was signed by 61 countries in January 1990 and went into effect in September 1990 (Lurie & Tjelflaat, 2012).

The document consists of a preamble and 54 articles. The preamble references other significant documents, such as the Universal Declaration of Human Rights, and notes that “the child, by reason of his physical and mental immaturity, needs special safeguards and care, including appropriate legal protection, before as well as after birth” (United Nations Office of the High Commissioner [OHCHR], 1989). While the UNCRC addresses a broad range of rights that should be guaranteed to all children by the signatory nations, the basic principles include nondiscrimination, the obligation to always act in the best interests of the child, promotion of survival and development and allowing young children to participate in decisions and older children to make their own decisions (Verhellen [The Convention on the Rights of the Child: Reflections from a Historical, Social Policy, and Educational Perspective], 2015). These rights have been characterized as the three Ps, including protection, provision (of certain basic goods and services) and participation, including the right to participate in decisions affecting the child (Verhellen [The Convention on the Rights of the Child: Reflections from a Historical, Social Policy, and Educational Perspective], 2015). The UNCRC is groundbreaking in that it establishes explicit public participation rights for children and, in doing so, overturns the dated Enlightenment notion that rights are intended for adults only (Wall, 2017).

By combining protection, provision and participation rights in a single document, the UNCRC moves children to the human rights agenda as legitimate public participants; children are now able to access public rights directly (Wall, 2017). The UNCRC was eventually signed by every other nation in the world. The United States remains the only holdout country in the world which has not ratified the treaty, despite the fact that President Obama signed it while in office. American critics of the UNCRC tend to be conservative and believe that the UNCRC grants children rights which could undermine traditional family values; which would allow American teens to consent to health care, including abortions; which would prohibit corporal punishment in schools; and which would prohibit capital punishment and solitary confinement of minors in prison. American supporters of the UNCRC note that American law trumps ratified treaties and controversial UNCRC provisions would not necessarily change current American practice or law.

Author Mary John (2003) has suggested that a fourth P, power, needs to be added to the three Ps noted above if the UNCRC is going to have a meaningful impact on the lives of children. She argues (2003) that "[i]n order to learn about power, children need to be given opportunities to exercise it" (p. 48). In an ideal setting, children will use their participation rights and, over time, become more skilled and empowered until they reach adulthood. Unless children are allowed to utilize this right, they are at risk of, at best, of being viewed as second-class citizens with limited rights or, at worst, of having voices which go unheard. Parents and others are being asked to change their vision of childhood to one that balances the right to participate in a modern world while still safe from those who might endanger vulnerable children.

The best tool to allow their voices to be heard may be digital tools given their ubiquity, but the UNCRC makes no explicit reference to them. Livingstone (2014), in a valuable contribution, has suggested that participation rights enunciated in the UNCRC might be updated today to be read to include the digital rights of "enhanced connections and networking opportunities, scalable ways of consulting children about governance user-friendly forums for child/youth voice and expression, child-led initiatives for local and global change, peer-to-peer connections for entertainment and learning and recognition of child/youth rights responsibilities and engagement" (p. 23). She infers these rights from Article 3 of the UNCRC, which references the child's participation in administrative and legislative proceedings if it is in the child's best interests. Similarly, Articles 12 and 13 grant children the right to express their opinions and to freedom of expression within the limits imposed by the law.

Article 12 states

1. States Parties shall assure to the child who is capable of forming his or her own view the right to express those views freely in all matters affecting the child, the views of the child being given due weight in accordance with the age and maturity of the child.

2. For this purpose, the child shall in particular be provided the opportunity to be heard in any judicial and administrative proceedings affecting the child, either directly, or through a representative or an *appropriate body, in a manner consistent with the procedural rules of national law (United Nations Office of the High Commissioner [OHCHR], 1989).*

Importantly, children's rights identified in Article 12 are not discretionary; parents and others do not have the option of ignoring a child's wishes. Children should be presumed competent (as opposed to being presumed incompetent, the routine paradigm in medical and legal proceedings today) and should be informed exactly how their views have been considered. Finally, non-verbal communication (of both disabled and—presumably—non-verbal children, should be recognized as expressing a view (Freeman, 2020, citing CRC Committee General Comment No. 12).

Article 3 notes that "[i]n all actions concerning children, whether undertaken by public or private social welfare institutions, courts of law, administrative authorities or legislative bodies, the best interests of the child shall be a primary consideration." Freeman (2020) writes that even governmental decisions such as building new roads, going to war and restructuring a school syllabus would presumably be governed by Article 3. While this may appear impractical on its face, the reasons are significant. Chief among them is the fact that children are inherently vulnerable and "[i]n a world run by adults, there is a danger that children's interests would otherwise be overlooked" (Freeman, 2020, p. 99).

Yet we know that children's interests **are** routinely overlooked. Furthermore, the prospect of children, including perhaps very young children, using digital media raises a number of challenges. Perhaps most fundamentally, we know that children around the world do not all share equal access to digital media, which limits their voices in all discussions—local, national and international—where their input is needed. Even when children have access, the issue of "evolving capacities" noted in the UNCRC imposes some limits on the use of media and other tools. Parents and other adults will need to safeguard their children to ensure that their rights are protected until children exhibit the requisite skills to protect themselves. How do we know when children have reached this stage? As we will see, even medical providers are not always sure when this stage is reached. What if only mobile devices are available? Does this impose any special obligations on parents and child advocates? Are they obligated to serve as good models, perhaps as both responsible digital media users and good digital citizens, to their children? Finally, what guidance has the international community provided through legislation, policies and other guidance?

GLOBAL DIGITAL OPPORTUNITIES FOR CHILDREN

Digital media are an interesting phenomenon in that they both pose unparalleled opportunities and significant threats to children. Parents, child advocates and legislators must balance the opportunities for growth presented with possible risks. First, access is important, but access alone does not ensure that children will necessarily possess the requisite skills to negotiate the internet. In the United States, "access to media technology does not guarantee access to the forms of capital—social and cultural—that are the crucial gateway to educational achievement, economic development, and political engagement" (Watkins [Preface], 2018a, xi). Digital tools, after all, are simply tools in the hands of the user. It is vital that schools and societies invest in both in preparing the user and the infrastructure to support digital media. "Computational thinking, critical thinking, and expert thinking, for example, are vital assets in our innovation economy" (Watkins [Preface], 2018a, xi). This suggests that investments in schools, curricula and students will also be an important part of our preparations for the future.

Yet despite even significant investment, opportunities for students are not evenly distributed throughout the United States. One of the greatest challenges American schools face is that "black, Latino and lower-income students…often lack access to the instructional expertise and curricula resources that develop the cognitive skills that drive our knowledge-based economy" (Watkins, 2018a, p. xi). In the U.S., homes with broadband access have overwhelmingly tended to be white or Asian, higher-income and higher-educated individuals (Watkins, 2018c, p. 57). This disparity, one of many in the United States, becomes crucial since so many vital skills are linked to competence in media skills. As identified by Jenkins (2006), these include play (problem-solving), performance (improvisation and discovery), simulation (interpreting and constructing dynamic models of the real world), appropriation (sampling and remixing media content), multitasking, distributed cognition (interacting meaningfully with tools able to expand mental capacities), collective intelligence (pooling knowledge and comparing notes with others to achieve a common goal), judgment (evaluating credibility of sources), transmedia navigation (following stories across different modalities), networking (synthesizing information) and negotiation (traversing diverse communities and respecting different perspectives (*Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*). The ability to meaningful use media in tandem with key cognitive skills is crucial if children are to thrive in our collective digital future.

By focusing less on the amount of time children spent on the internet and more on their repertoire of skills, parents can ensure that children are developing the skills they need to succeed in school and later in careers. Studies have shown, for instance, that gaming can teach the skills that are needed for successful learning (Watkins [How Black and Latino Youth are Remaking the Digital Divide], 2018b, 44-45). The reality is that student access, internet competency and skills and success in school are complex and intertwined issues which vary considerably from community to community.

While the access divide in the United States and other first world countries has become less pronounced in recent years, measuring access is more nuanced. It has been suggested that society should not use binary measures (inclusion/exclusion) to ascertain whether children are digitally active since digital skills should be viewed as capacities which enable a child to enhance his/her/their future (Third, et al, 2019). This approach suggests that access is merely the first step in a long process in which parents, teachers and schools work together to ensure that children develop the requisite skills to thrive in a digital world. While the first digital divide involves those with and without access, the second digital divide is skills-based (Lindgren, 2017).

Even when access is not an issue, however, children may not understand fully the rules governing internet use. A 2017 report by England's Children's Commission entitled *Growing Up Digital* attempted to measure children's understanding of the 5,000-word, 17-page Terms and Conditions of Instagram. After 20 minutes of reading, the 13-year-olds had only gotten halfway through the document and many asked to stop. Importantly, this document explains Instagram's right to track users even when it is not in use, buy and sell personal data, change terms without notice and terminate one's account with no notice (2017). If children are unwilling to spend the necessary time to read this document then, like many adults, they may not understand the risks they are voluntarily accepting.

Unlike developed nations, in developing countries, access is primarily mobile in nature. While in the U.S. there is some evidence indicating that homes with mobile-only internet access may be perpetuating disparities (Watkins [The Mobile Paradox], 2018c, 76), the outlook may be more optimistic overseas where mobile phones are often the norm. Mobile phones have recently become a vehicle of empowerment, especially for those who were long without a voice, such as sub-Saharan African young girls and women; phones can used to extend their education and network of contacts, develop their entrepreneurial skills and reach out to potential romantic partners (Porter et al, 2020). Where physical books may be difficult to obtain, phones can be used for educational and other purposes, according to a 2014 UNESCO report examining global cell phone use and facility (*Growing Up Digital).* This is especially important when one considers that 77 percent of European households have access to the internet, while only 7 percent of African households have internet access (*Growing Up Digital*, 2014). Interestingly, the study suggests that once girls and women start reading on their phones, they remain much more engaged than their male counterparts (*Growing Up Digital*, 2014).

Life in developed countries where primary digital access is not mobile provides some safeguards that their youth mobile-only counterparts overseas may not possess. What are the differences in the internet activities in which children engage which occur at home on a home computer compared to those which occur on a mobile phone? First, parents or other guardians may be able to play a protective role at home, especially if the computer is a shared one in a common space. Knowing that parents are likely to share the same space at home could prevent the child from visiting sites that could pose a danger to an unsophisticated child. Second, a public space at home could also discourage the child from visiting sites such as pornographic or dating sites. Third, parents have the ability to install mechanisms such as tracking devices to allow them to see which sites their child has visited and blocking devices or filters to prevent them visiting certain types of internet sites.

In 2015, U.S. children were reported to receive their first cell phone at age 6 (Aiken, 2016). Furthermore, 69 percent of young people reported that they hid at least a portion of their on-line activity from their parents (Aiken, 2016). It may be easier to hide activity on a mobile phone than on a shared home computer. For much of the world, cell phone use has become a cultural rite of passage.

Understanding “Evolving Capacities” of Children

As noted earlier, Article 12 of the UNCRC emphasizes the importance of allowing children who can form and express opinions to be heard in all matters affecting them, especially judicial and administrative issues which will impact them. Furthermore, “due weight” must be accorded to the child’s wishes. Articles 5 and 14 note that the “evolving capacities” of the child must be respected.

While children may not have equitable access to digital media in some parts of the world and may have access only to mobile phones in other parts, when they do access internet content, how can we be sure that they understand what they are viewing and should, in fact, have the right to access? That is, when does participation trump protection? The “evolving capacities” standard provides some guidance in this context.

The Committee on the Rights of the Child (CRC) noted that one of the three primary functions of the important term “evolving capacities” was as an interpretative principle, including Articles 3 (best interests of the child), Article 12 (right to be heard), Article 13 (freedom of expression), Article 14 (freedom of thought, conscience and religion) and Article 17 (right to information) (Varadan, 2019). Therefore, “parents and guardians no longer carry a *carte blanche* in how they provide guidance and direction to their children…as a child’s capacities evolve, increasing weight should be accorded to his or her views in the determination of best interests” (Varadan, 2019, p. 320-21).

This approach is not unlike that taken in the U.S. when minors exhibit mature comprehension of their medical condition in healthcare settings, but may not meet the chronological threshold for adult consent. Providers may use a sliding scale of competence to ascertain whether the minor understands the risks, benefits and other relevant information about a proposed medical procedure. The closer the child is to actual chronological adulthood, the more likely the physician is to defer to the child and the less likely the physician is to defer to the parents if the parents happen to disagree with the child's treatment preference. The younger the child—and therefore the less likely the child is to understand the short-term and long-term consequences of any medical decision—the less likely the provider would be to defer to parents. Physicians and parents can make better medical decisions for their pediatric patients when they (at a minimum), know their patient's preferences.

As a policy principle, "evolving capacities" unshackles States from traditional policy-making frameworks, in which children are presumed to lack capacity until they cross a specific age-barrier or reach a prescribed age of adulthood. It debunks the notion that children must reach a requisite threshold of capacity to be able to exercise their rights; and it recognises that as children grow and develop they need to be progressively enabled and empowered in the exercise of their rights" (Veradan, 2019, p. 329).

In the 2018 General Data Protection Regulation (GDPR), considered the strictest law of its kind, harsh fines are imposed on those who breach data privacy and other violations for European Union (EU) residents, regardless of where the organization is based in the world. One of its seven principles is consent. However, EU nations could not arrive at consensus about the age at which a minor should be able to consent to internet activity and privacy questions. Signing countries found that consent ranged from age 13 (Belgium, Denmark, Estonia, Finland, Latvia, Poland, Portugal, Spain, Sweden and U.K.), 14 (Austria, Bulgaria, Cyprus and Italy), 15 (Czech Republic, Greece and Slovenia), 15/16 (France) and 16 (Croatia, Germany, Hungary, Ireland, Lithuania, Luxemburg, Malta, Romania, Slovakia and the Netherlands) (Livingstone, 2018). Countries advocating for younger ages of consent would likely agree that participation rights trump protection rights, while countries employing older ages of consent would advocate protectionism over participation.

Interestingly, when parents in the U.K. were asked when they thought their child would be mature enough to visit websites or use apps, the average age chosen was 13, but the most common answer (the mode) was 16. Strikingly, parents of very young children considered 13 an appropriate age, while parents of teenagers thought the child should be at least 14 and 15 would be ideal (Livingstone, 2018). This suggests that parents who live with—and presumably know the cognitive strengths and weaknesses of their children—prefer the slightly older standard.

Judges, attorneys, social workers and others who question children in courtrooms, hospitals or similar settings, can make better decisions when they take into account the child's wishes. This does not necessarily mean defer to the child's wishes, however. This approach provides flexibility for those involved in making decisions that will impact the child since they are no longer bound by rigid age-dependent guidelines.

This approach is consistent with the view that children today are no longer in need of "child saving," but of "child empowerment" (Wall, 2017, p. 4). Learning to use and exercise power is an important skill where competence is only earned after practice. Much like assent of minors or informed consent in the medical treatment process, "[i]n order to learn about power, children need to be given opportunities to exercise it" (John, 2003, p. 98). As Vickery (2017) notes, we need to stop "conflating risk and harm. That students may encounter inappropriate content does not necessarily mean that they will be harmed by it" (p. 108).

Since the traditional cognitive development theory of Piaget that children mature in a series of discrete states is now viewed as inadequate to explain the complex cognitive maturation process, John

(2003) writes that giving rights to children incrementally based on chronological age is also a dated concept since it can be used to exclude some children (p. 63). Cognitive maturity varies a great deal and a blanket approach to rights-giving is not a respectful approach to this complex problem. Case-by-case determinations are likely called for.

The goal of determining cognitive maturity is to attempt to measure at what point on a scale from infancy to adulthood the child falls. The closer the child is to adulthood, the greater the child's participation rights and the greater the child's "evolving capacities" to participate in the digital community.

The reality is that for children just as for adults, participation rights are ultimately the key to gaining improved rights overall. It is true in part, again, that participation rights rely on other kinds of rights, such as protection against discrimination and provisions of education and resources. But without a renewed focus on children's participation rights, such as to voices, agency, and recognition in the public world, children will never escape their historical status as second-class citizens. (Wall, 2017, p. 160).

Efforts to Teach Children to Be Digitally Resilient

The development of skills necessary to negotiate the internet and learn to calculate risk allows children to develop what has been called digital resilience. "The 'ideal' digitally resilient young person thus emerges at the centre of dominant framings of resilience as someone who both prevents and reacts to online risks by exercising high levels of awareness; technical, cognitive and communicative skills; and strategies of self-regulation" (Third, et al, 2019, p. 63). Two London School of Economics researchers found that parents who attempted to shield their children from on-line content the parents found offensive tended to have children who were less resilient than parents with children who were not routinely protected from internet content (Aiken, 2016). Just as those who survive catastrophic events such as 9/11 in the United States eventually become stronger, children who learn to either avoid or leave sites not intended for children learn the lesson and become digitally stronger and wiser. In effect, one unfortunate visit to an inappropriate site (due to content or some other reason) instructs that child not to make the same error again. The next visit to the internet will be more informed.

One possible advantage of visits to the internet by a more informed child is an increased understanding of the role of the digital citizen. Many countries, especially Western nations, have developed policies explicitly designed to protect—or even control--children online, even when they are exercising citizenship-type rights similar to those of adults. For instance, the Australian Communications and Media Authority (ACMA) created a three-pronged approach to internet access in 2009 emphasizing digital etiquette, digital literacy and digital security. However, in 2013, the same group created new guidelines focusing on 'positive engagement' and 'being cybersmart' (Third, et al, 2019, p. 184-85). In just a few words, Australia effectively switched its policy approach from one of child protection to child participation. Participation allows for children to enhance the requisite skills necessary to become global citizens.

For many young people, digital citizenship is routine. Unlike their parents, the internet and mobile phones are the way they routinely communicate with their school friends and with friends around the world. This is what has been labelled the "digital everyday" for youth today: "digital micropractices of citizenship" (which includes fandom, special interest communities, sharing political content on social media, signing petitions, blogging, etc.) (Third, 2019, p. 181). These micropractices have become a socialization tool for young people today.

There is evidence to show that socialization and digital citizenship has meaningful consequences in the lives of children. The 2019 International Fund for Agricultural Development (IFAD) *Rural Youth Inclusion, Empowerment and Participation* report reviewed 54 programs focused on youth participation in southern Asia, sub-Saharan Africa and Latin America to consider their roles in informing, consulting, collaborating and empowering (Trivelli & Morel, 2019). Interesting current projects include a Sri Lankan program designed to strengthen communication technology skills among youth, an African Union Continental Youth Consultation on Transitional Justice designed to solicit the opinions of young people on diversity and civil rights issues and youth parliaments in 14 countries (Trivelli & Morel, 2019). The authors suggest that “when thinking about youth inclusion, participation is decision-making” (Trivelli & Morel, 2019). Unless the interaction has some meaningful outcome, the child is not empowered.

The Children’s Commissioner for England reexamined the UNCRC and the issue of digital rights of children and affirms the notion that, while digital tools are now widely available, the impact of children’s voices is not as significant as child advocates had hoped.

[T]he greater availability of digital media is not being used to include or amplify children’s voices in the design of interventions and decision-making processes, with considerable digital and cultural barriers to children being heard and responded to…child participation, even in an age of digital connectivity, is still more promise than reality, and both determination and guidance from states are sorely needed, especially given the considerable attention to risk-focused and protectionist—sometimes overly protectionist—approaches to digital media (Livingstone, et al, 2017, p. 16).

Children have identified the right to access information as the most important right impacted by technology (Livingstone, et al, 2017, p. 18). However, policies designed to minimize content risks and access can result in other concerns, such as misinformation involving health issues, which could lead to harmful health practices (Pascoe, 2011). Adults and others hoping to protect children from specific information may, in fact, be posing a danger depending on the strategy used.

DIGITAL MEDIA POLICY RECOMMENDATIONS

First, the sole holdout to ratification of the UNCRC, the United States, should join the rest of the world in signing this document that advances the agenda of children everywhere. While many provisions remain aspirational, this should not prevent the U.S. from—at a minimum—starting the necessary work of ensuring that the rights of children are protected. While it is certain that some signing countries have given little or no thought to the UNCRC and its implications for the children of their respective nations, the United States nevertheless loses credibility when attempting to create child-friendly laws and policies without having previously ratified the UNCRC. Furthermore, given the likely objections of some American politicians to specific provisions of the UNCRC, it is possible to sign on to an amended version of the treaty (Wall, 2017).

Second, all schools should add a digital citizenship component to their curricula. Starting in grade school, children should learn the tools and skills necessary to become competent at negotiating the internet. At a minimum, these skills should include reducing cyberbullying, ensuring that younger children understand that they will be using computers equipped with filters, firewalls and other devices designed to protect them, understanding that they have the right to privacy and confidentiality (which may, at times, need to give way to parental authority in the case of younger children), access to information for education and entertainment purposes, an understanding of digital citizenship and their right to be heard and speak up on issues of importance to them and an understanding of the role that the rights granted to them by the UNCRC. As the child matures, their "evolving capacities" would mandate that increasing deference would be granted to the child and less to the parents. The child should be granted the opportunity to exercise and develop these rights.

Third, an app or other tool designed to verify age prior to accessing specific internet sites would address parental safety concerns for very young children. This might be a challenging task as it could conceivably be easy to breach if older children or siblings were willing to share passwords, but this should not deter parents seeking to limit access to certain websites that they do not want their children to access. Parents or guardians could determine which sites they deemed appropriate for their children on a case-by-case basis.

Fourth, all signing countries should endeavor to make the provisions of the UNCRC known to their citizens so that all residents, including children, are aware of their rights. For children, the information can be provided in the school curriculum noted earlier. However, to become part of the societal fabric of each country, this information involving fundamental rights of children needs to widely disseminated and, in some case, may involve a fundamental values shift. After all, notorious human rights abusers such as Sudan, the Democratic Republic of the Congo and China have all signed the document, yet are not known to protect human rights, including children's rights (Blanchfield [The United States Convention on the Rights of the Child: Background and Policy Issues], 2011).

In order to achieve this goal, those countries which have been most successful at incorporating the values of the UNCRC into their societies have done so by changing their Constitution or creating legislation designed to achieve UNCRC goals. In a UNICEF study, nations which demonstrated high levels of commitment to UNCRC principles as evidenced through practices such as including them in their respective Constitutions or creating new legislation were places where "children were perceived as rights-holders and that there was a culture of respect for children's rights" (Lundy, et al, 2013, p. 451). Countries which chose to incorporate the principles in their Constitutions or in other public policy ways, established vital leverage for politicians and NGOs to use in order to ensure integration of the UNCRC principles into national law and policy (Lundy, et al, 2013). Since these approaches have proven more successful at achieving UNCRC goals, nations should consider the (admittedly politically challenging) approach of addressing Constitutional or legislation changes; the advantage appears to be a climate which is more UNCRC-friendly.

Fifth, nations should adopt terms and conditions for apps and digital media planforms which are both shorter and use more child-friendly language. Children (and indeed, many adults) may not realize the long-term implications of the documents they routinely sign online. By emphasizing the consequences of how their personal information will be used in the future, children will, at a minimum, be able to make more informed decisions.

DIGITAL QUESTIONS FOR THE FUTURE

Much of this chapter has focused on participation and, more meaningfully, empowerment and how to ensure that children have the requisite tools to move beyond participation so that they are empowered to make decisions which have meaningful consequences. Yet, there is little in the literature to measure empowerment. What does child empowerment look like? Furthermore, it will likely look very different in the United States compared to a developing country. One area of research that needs to be explored is the creation of tools to measure empowerment. What unit of measurement should we use? By hours per week? By discrete activity? Are all online activities equally valuable?

Also, there is evidence indicating that specific vulnerable children, especially girls, children living in rural areas and those with disabilities, may benefit more than others by empowerment through the use of digital tools and access to information and by communicating with other children (Third, et al, 2019). A more thorough exploration of which specific groups of children and which internet sites boost empowerment would be beneficial to communities seeking to enhance child rights.

Since digital access has been shown to be so important to all peoples, not just children, governments should consider investing in technologies designed to make Wi-Fi widely available. Since so many developing countries still rely exclusively on mobile phones for internet access and internet access has been shown to be essential for multiple purposes, governments have an obligation to ease access to the extent of their financial ability. Where possible, nations should seek to make Wi-Fi widely available, including in underserved and rural areas.

The notion of resilience has been considered often in the context of emergencies and survivorship after tragedies such as war and 9/11. The implications of digital resilience are not yet clear. Unanswered questions include whether digital resilience is measurable and how best to create curricula designed to ensure that young people develop these skills. What characteristics do we want to see emerge in a digitally-resilient child? How will we ensure that these skills remain up-to-date in adults? Should we automatically assume that adults are digitally resilient?

Furthermore, unanswered questions remain about how best to both measure and determine when children have the cognitive skills necessary to bypass internet information which is clearly not reliable. How can we ascertain when children possess the requisite cognitive skills to avoid specific content and search for more reliable content? What sort of curriculum is required to pass those skills on to children? In a related question, is it possible to research the issue of whether the UNCRC policy goals of improving the lives of children successfully move from theory to practice by enhancing their digital lives?

CHILDREN AND THE FUTURE OF GLOBAL DIGITAL MEDIA

What was once a luxury for many people—digital access—is now necessary for basic communication, work, entertainment and other purposes. In order to meet the needs of children around the world today, parents, legislators and those who simply care about the future of children, need to ensure that children have the tools to allow their voices to be heard. Without the requisite infrastructure, digital access and government support, children run the risk of not being heard in decisions affecting their individual and collective futures. While the internet is generally viewed as a democratization mechanism capable of leveling playing fields rendered uneven by health, wealth and other asset distribution inequities, adults have an obligation to remove obstacles when possible. Unless the voices of those traditionally impacted by these inequities are allowed to be heard, however, the lofty goals imagined by the UNCRC will not be achieved.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Aiken, M. (2016). *The Cyber Effect: A Pioneering Cyberpsychologist Explains how Human Behavior Changes Online*. Random House.

Blanchfield, L. (2011). The United States Convention on the Rights of the Child: Background and Policy Issues. In K. McGowan (Ed.), United Nations: International Agreements and Efforts. Academic Press.

Children's Commissioner's Growing Up Digital Taskforce. (2017). *Growing Up Digital: A Report of the Growing Up Digital Task Force*. Children's Commissioner. https://www.childrenscommissioner.gov.uk/report/growing-up-digital/

Convention on the Rights of the Child. Nov. 20, 1989, 1577 U.N.T.S. 3. www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx

Freeman, M. (2020). *A Magna Carta for Children? Rethinking Children's Rights*. Cambridge University Press.

General Data Protection Regulations. (2018). https://gdpr-info.eu

Jenkins, H. (2006). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. MacArthur Foundation.

John, M. (2003). *Children's Rights and Power: Charging Up for a New Century*. Jessica Kingsley Publishers.

Lindgren, S. (2017). *Digital Media and Society*. Sage.

Livingstone, S. (2014). Children's Digital Rights: A Priority. *LSE Research Online.*, *42*(4/5), 20–24.

Livingstone, S. (2018). Children: A Special Case for Privacy? *Intermedia.*, *46*(2), 18–23.

Livingstone, S., Lansdown, G., & Third, A. (2017). *The Case for a UNCRC General Comment on Children's Rights and Digital Media.* LSE Consulting. https://www.childrenscommissioner.gov.uk/report/the-case-for-a-uncrc-general-comment-on-childrens-rights-and-digital-media/

Lundy, L., Kilkelly, U., & Byrne, B. (2013). Incorporation of the United Nations Convention on the Rights of the Child in Law: A Comparative Review. *International Journal of Children's Rights*, *21*(3), 442–463. doi:10.1163/15718182-55680028

Lurie, J., & Tjelflaat, T. (2012). Children's Rights and the UN Convention on the Rights of the Child: Monitoring in Norway. *Dialogue in Praxis, 1*(14), 41-56.

Pascoe, C. (2011). Resource and Risk: Youth Sexuality and New Media Use. *Sexuality Research & Social Policy*, *8*(1), 5–17. doi:10.100713178-011-0042-5

Porter, G., Hampshire, K., Abane, A., Muthali, A., Robson, E., De Lannoy A., Tanle, A. & Owusu, A. (2020). Mobile Phones, Gender, and Female Empowerment in Sub-Saharan Africa: Studies with African Youth. *Information Technology for Development, 2*6(1), 180-93.

Third, A., Collin, P., Walsh, L., & Black, R. (2019). *Young People in Digital Society: Control Shift*. Palgrave MacMillan. doi:10.1057/978-1-137-57369-8

Trevelli, C., & Morel, J. (2019). *Rural Youth Inclusion, Empowerment and Participation* (Report No. 45). International Fund for Agricultural Development. www.ifad.org/ruraldevelopmentreport

United Nations Convention on the Rights of the Child. (1989).

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2014). *Reading in the Mobile Era: A Study of Mobile Reading in Developing Countries*. https://unesdoc.unesco.org/ark:/48223/pf0000227436

Varadan, S. (2019). The Principle of Evolving Capacities under the UN Convention on the Rights of the Child. *International Journal of Children's Rights*, *27*(2), 306–338. doi:10.1163/15718182-02702006

Verhellen, E. (2015). The Convention on the Rights of the Child: Reflections from a Historical, Social Policy and Educational Perspective. In W. Vanderhole, E. Desmet, D. Raynaert, & S. Lambrechts (Eds.), *Routledge International Handbook of Children's Rights Studies* (pp. 43–59). Routledge.

Wall, J. (2017). *Children's Rights: Today's Global Challenge*. Rowman & Littlefield.

Watkins, C. (2018a). Preface. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Divide: How Black and Latino Youth Navigate Digital Inequality* (pp. ix–xiii). NYU Press.

Watkins, C. (2018b). How Black and Latino Youth are Remaking the Digital Divide. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Edge: How Black and Latino Youth Navigate Digital Inequality* (pp. 19–49). NYU Press.

Watkins, C. (2018c). The Mobile Paradox: Understanding the Mobile Lives of Latino and Black Youth. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Edge: How Black and Latino Youth Navigate Digital Inequality* (pp. 50–77). NYU Press.

ADDITIONAL READING

Alaimo, K., & Klug, B. (Eds.). (2002). *Children as Equals: Exploring the Rights of the Child*. University Press of America.

Baruch, A., & Erstad, O. (2018). Upbringing in a Digital World: Opportunities and Possibilities. *Technology, Knowledge and Learning.*, *23*(3), 377–390. doi:10.100710758-018-9386-8

Byrne, B., & Lundy, L. (2013). Reconciling Children's Policy and Children's Rights: Barriers to Effective Government Delivery. *Children & Society*, *29*(4), 266–276. doi:10.1111/chso.12045

Freeman, M. (2014). *The Future of Children's Rights*. Brill.

Goldhagen, J., Shenoda, S., Oberg, C., Mercer, R., Kadir, A., Raman, S., Waterston, T., & Spencer, N. (2020). Rights, Justice and Equity: A Global Agenda for Child Health and Wellbeing. *The Lancet. Child & Adolescent Health*, *4*(1), 80–90. doi:10.1016/S2352-4642(19)30346-3 PMID:31757760

Rutgers, C. (Ed.). (2011). *Creating a World Fit for Children: Understanding the UN Convention on the Rights of the Child.* New York: International Debate Education Association.

Schulz, W., & Raman, S. (2020). *The Coming Good Society: Why New Realities Demand New Rights*. Harvard University Press. doi:10.4159/9780674245792

Vandenhole, W., Desmet, E., Raynaert, D., & Lembrechts, S. (Eds.). (2015). *Routledge International Handbook of Children's Rights Studies*. Routledge. doi:10.4324/9781315769530

Vickery, J. R. (2017). *Worried About the Wrong Things: Youth, Risk, and Opportunity in the Digital World*. MIT Press. doi:10.7551/mitpress/10653.001.0001

KEY TERMS AND DEFINITIONS

Adolescent: Young person under 18 years of age.

Child Empowerment: Recognition that children have traditionally been denied the opportunity to participate in decisions impacting their lives and attempt to correct historical incidents where children's wishes where either routine ignored or their opinions were not obtained.

Cognitive Maturity: The process of increasing intellectual reasoning and ability whereby knowledge is acquired typically associated with increasing chronological age prior to adulthood.

Digital Citizenship: Role that children can play online as citizens of the world; examples include minor activities such as circulating and signing petitions or significant roles such as encouraging nations to change their political agendas or adopt specific political platforms.

Digital Resilience: Concept that exposure to potentially-disturbing or age-inappropriate internet content (e.g., pornography) by children will result in children who are able to bounce back quickly and recover. This skill allows users to be more prepared to encounter inappropriate content in future internet visits and is considered an important part of the growth process by many who argue that participation should trump protection.

Evolving Capacities: The notion that parents, physicians, social workers, judges and others who work with children should ascertain their level of understanding of any proceeding that will have an impact on the life of the child and seek their input in the decision-making process. Generally, the higher the level of understanding of the child, the more input the child will have in the process.

United Nations Convention on the Rights of the Child (UNCRC): Treaty ratified by every country except the United States which grants children unparalleled rights in virtually all aspects of their lives.

Chapter 5
Instructing AI Ethics and Human Rights

Katharina Miller
https://orcid.org/0000-0002-2088-4531
European Women Lawyers Association, Spain

Muhammet Demirbilek
Faculty of Education, Suleyman Demirel University, Turkey

ABSTRACT

This chapter will exploit emerging issues of AI and current literature on AI ethics and human rights teaching. The authors will exploit understanding AI ethics and human rights in daily life and offer teaching methodologies to explain how to teach AI ethics and human rights in K-12 learning environments. Furthermore, the chapter will be devoted to the latest trends and issues on how to teach AI ethics and humans rights teaching in K-12 learning environments. Particular emphasis will be made on a survey of existing ethics teaching methodologies and how to adopt existing teaching strategies into AI ethics teaching in order to improve their understanding on AI ethics and human rights.

INTRODUCTION

Industrial and governmental organizations are harnessing the power of AI algorithms and applications that are already revolutionizing a wide range of economic sectors that have already started to impact our daily lives including education, marketing, transportation, communication, finance, and customer services.

Over the past decade, AI technology has progressed exponentially, becoming a key fabric of our everyday lives across e-commerce, educational and research platforms, entertainment, and the popular imaginary.

While marvelous advancements in AI technology have progressed exponentially, the potential misuse of AI has become main concern. There is a high demand to establish ethical standards and regulations for the daily use of AI. Furthermore, there is a need to teach AI ethics and human rights in schools to

DOI: 10.4018/978-1-7998-8467-5.ch005

explore AI technology and human rights from ethical perspective and to increase students' familiarity with ethical use and implications and impact on of these technologies on human rights.

This chapter will exploit emerging issues of AI and current state of art literature on AI ethics and human right teaching.

The authors will exploit understanding AI ethics and human rights in daily life and offer teaching methodologies to explain how to teach AI ethics and human rights in K-12 learning environments. Furthermore, the chapter will be devoted to the latest trends and issues on how to teach AI ethics and humans rights teaching in K-12 learning environments.

Particular emphasis will be made on the survey of existing ethics teaching methodologies and how to adopt existing teaching strategies into AI ethics teaching in order to improve their understanding on AI ethics and human rights.

INTRODUCTION TO ETHICS AND HUMAN RIGHTS

Ethics History and Definitions

Philosophical ethics is a discipline within philosophy and it can be understood as the science of moral action. It examines human practice in terms of the conditions of its morality and attempts to establish the concept of morality. In this context, morality means the quality that allows an action to be described as moral, as a morally good action (Pieper, 1991). Therefore, ethics can be described as "practical" philosophy (Vieth, 2006). Already Immanuel Kant (1785) stated that not only philosophers or ethicist can discuss moral questions:

"Common sense can just as well hope to get it right as a philosopher can always hope to get it right; indeed, it is almost more certain of it than even the latter, because the latter, however, has no other principle than the former, but can easily confuse his (or her) judgment by a multitude of strange, irrelevant considerations and make it deviate from the straight direction (…). There is something splendid about innocence; but what is bad about it, in turn, is that it cannot protect itself very well and is easily seduced. That is why even wisdom - which otherwise probably consists more in doing things than in knowledge - also requires science, not in order to learn from it, but to give its prescriptions entrance and permanence."

Kant thus assumes that the layperson, i.e., a person who is not explicitly philosophically educated, is more reliable than the philosopher in finding the right thing to do.

There are many different fields within the philosophical ethics such as fundamental ethics, applied ethics, feminist ethics, etc.

The solution of ethical problems in applied ethics is not only to enforce philosophically satisfying arguments and concepts, but to reflect the evaluative experience of persons involved in a context experienced as problematic. Philosophical arguments and concepts can and must correct the evaluative experience of persons, but may only do so cautiously. One speaks of evaluative experience because the way in which people are confronted with situations brings value aspects to light and these values can vary depending on the context of each person (his or her nationality, religion, sex, age etc.). It is about problem solving in concrete situations and that means essentially that Applied Ethics conveys practical orientation to the people involved. In this chapter, the authors will treat AI Ethics as part of Applied Ethics. AI Ethics will be treated as doubly normative: on the one hand, the authors presuppose that ethics as a *theory of right and good* must be compatible with the *value of pluralism*. On the other hand, one must reflect the individual practical experience of people, even if one disagrees with them. Another important aspect of (applied ethics) is a statement by a German sociologist who stated: "Ethics is a "bicycle brake on an intercontinental airplane" (Beck, 1988). Beck basically expressed the following concerns: in his opinion, factual plurality of philosophical points of views leads to the end of the ethical commitment. Furthermore, statements by philosophers on medical, legal, technical, biological and other issues are presumptuous because philosophers only have non-professional competence. Two arguments can be brought up against Beck's statements: Factual plurality of philosophical points of view does not lead to the end of the ethical commitment. Pluralism is itself an ethical value, which becomes relevant precisely because of the importance of autonomy in applied ethics in democratic societies. Applied Ethics is not about discovering a universal set of obligations. AI Ethics is not primarily about ethical (universal) obligations, but rather about practical orientation. One of main concerns is to eliminate irritations in the evaluative experience of people in situations, such as having sex robots. For some, these kinds of robots are just a normal thing to have, whereas many people, and especially male and female feminists, worry about the future of human sexual relations. Answers to these questions are not given by ethicists, but are developed by the persons concerned in dealing with situations. Doing so, the greatest enemy of ethics is the certainty that you have always done it right or that it is good the way it is. Coming back to the example of sex robots: the argument that only because sex robots have not existed in the past (which is not totally true, if one follows the arguments by Adrienne Mayor (2018) who sees Pasiphaë's hollow wooden cow created by Daedalus, and which she climbed into in order to mate with the bull as a sextech object) they should not be created nor used.

A descriptive understanding of AI Ethics does not recognize normative reasons, but only factual causes for the validity of certain standards. The question of what the right moral is, is considered pointless. Descriptive ethics therefore does not ask for ethical reasons for the applicable prohibitions and does not criticize the individual way of life

One looks at a society from the point of view of an ethnologist or sociologist who describes carefully. Coming back to the sex robots: in a society the prohibition of sex robots then applies because the law enforcement agency does sanction the possession of sex robots (or sex toys in general).

A descriptive understanding of ethics is fundamentally senseless, insofar as it assumes that all questions of normative orientation are non-moral. it is senseless because AI Ethics already presupposes the meaningfulness of these questions.

A normative understanding of AI Ethics presupposes that it is possible to cite ethical reasons (value aspects) for the validity of standards. The question of what the right moral is considered meaningful. The prohibition of sex robots will be based on moral grounds derived from some countries current understanding of their religion or their values. Following these examples, it makes sense to distinguish between normative and descriptive aspects of justifications. Since there are standards or prohibitions not only in the ethical context, but also in the technical-economic (e.g., ISO standards, etc.) or legal area (laws, regulations, etc.), not all questions regarding the justification of standards touch AI Ethics. This does not mean, however, that all questions are ethically neutral in terms of technology, economy and law. The last important definition within this AI Ethics introduction is moral versus immoral: if something is assumed to be ethically relevant, then one can ask oneself whether it is morally good or morally bad or right or wrong. If, for example, the use of embryos as opposed to stones raises certain ethical-moral questions and something in the non-moral sense is an embryo, then some actions involving embryos are moral and some are immoral. This was at least the case for the Horizon 2020 Projects of European Union: Research involving Human Embryonic Stem Cells were not eligible for funding if they are directly derived from embryos within a Project. If they derived from previously established cells lines then a project used to be eligible for funding. If the research led to the destruction of Human Embryos, then the research was excluded. An AI Ethics related example could be use of Lethal autonomous weapon systems (LAWS). They can be missiles capable of selective targeting to learning machines with cognitive skills to decide whom, when and where to fight without human intervention. According to the Ethics Guidelines for Trustworthy Artificial Intelligence of the High-level expert group on artificial intelligence (AI HLEG 2019) these LAWS raise ethical-moral questions. The AI HLEG (2019) fears “an uncontrollable arms race on a historically unprecedented level, and create military contexts in which human control is almost entirely relinquished and the risks of malfunction are not addressed”.

Human Rights History and Definitions

According to the United Nations Human Rights Office of the High Commissioner, human rights are rights that every human has simply because he or she exists as human beings - they are not granted by any state. These universal rights are inherent to all human beings, regardless of nationality, sex, national or ethnic origin, colour, religion, language, or any other status. They range from the most fundamental - the right to life - to those that make life worth living, such as the rights to food, education, work, health, and liberty.

In 1948, human rights were established as fundamental and universally protectable in a legal document. This is the Universal Declaration of Human Rights (UDHR), which was adopted by the UN General Assembly. Now 73 years old, the UDHR remains the basis of all international human rights law. It consists of 30 articles that form the principles and building blocks of current and future human rights conventions, treaties and other legal instruments such as the Charter of Fundamental Rights of the European Union (EU) which was declared in 2000, and came into force in December 2009 or the European Convention on Human Rights and the European Social Charter

Together with the International Covenant for Civil and Political Rights, and the International Covenant for Economic, Social and Cultural Rights, the UDHR forms the International Bill of Rights.

The cornerstone of international human rights law is the principle of universality of human rights, which means that all human beings are all equally entitled to human rights. This principle is reiterated in many international human rights conventions, declarations and resolutions.

Human rights are inalienable. They should not be taken away except in very specific situations and only after due process. For example, the right to liberty can be restricted, as has happened in the Covid19 situation around the world. AI can threaten human rights such as the right to equality, the prohibition of discrimination and the right to privacy.

It is important to note that all human rights are indivisible and interdependent. One set of rights cannot be fully enjoyed without the other. For example, when civil liberties were restricted during the Covid pandemic, economic, social and cultural rights were affected. Similarly, i.e., AI based technologies violating social rights such as the right to equality or non-discrimination can have a negative impact on many other rights.

Individuals are entitled to human rights and they are obliged to also respect and stand up the human rights of others.

AI RELATED ETHICAL ISSUES

There is no universally accepted definition of AI (European Union Agency for Fundamental Rights, 2020), that's why the authors of this chapter refer to AI as per the (updated) definition by the High-level expert group on artificial intelligence (AI HLEG 2019):

"Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."

Some of the ethics issues related to AI are bias, fairness, safety, transparency and accountability (West and Allen, 2020). Not addressing or applying these ethics issues can lead to violation of human rights. There are many initiatives underway with regards to the ethical aspects of AI and one of them is the AI HLEG.

According to the AI HLEG (2019), AI related bias issues exist because "many AI systems, such as those including supervised machine learning components, rely on huge amounts of data to perform well (…)". The bias gets into the data because humans who create the datasets or set the instructions for the machine learning bring in her or his own bias, and this can happen consciously or unconsciously. Cathy O'Neil (2016) explains this perfectly with the example of a mother who wants to communicate her "internal dynamic cooking model" to a third person by formalizing this model, "making it much more systemic and, in some sense, mathematical". These kind of models "despite their reputation for impartiality, reflect goals and ideology", because human beings always feed models with their own values and desires. According to the AI HLEG and many other activists, "unfair bias must be avoided, as it could have multiple negative implications, from the marginalization of vulnerable groups, to the exacerbation of prejudice and discrimination." Especially the bias aspects are very relevant for AI Ethics as the evaluative experience is different for every human being; in an automized world, it's very difficult to maintain the evaluative experience. In the following some examples from within the hiring funnel shall illustrate in which ways AI is already interfering with people's very own evaluative experience. This was the case for Amazon's AI recruiting tool, which showed bias against women (Dastin, 2018). Amazon had tried to develop a tool that selects automatically the top five applicants out of hundreds. The responsible staff at Amazon trained their computer models to check applicants by observing patterns in CVs submitted to Amazon over a 10-year-period. At that time, most CVs had been submitted by men which reflected the lack of visibility of woman across the tech industry at that time (and which still exists). That's how Amazon's machine learning taught itself that male candidates were preferable over women. The tool penalized CVs that included the word "women's," as in "women's chess club captain." Furthermore, Amazon's AI recruiting tool downgraded graduates of two all-women's colleges because they did not specify the names of the schools (Dastin, 2018). It is as easy as "garbage in, garbage out" (GIGO), meaning that the specific data you fill into a machine comes out as this specific data.

Another recruiting tool claims to measure the performance of applicants in video interviews by automatically analyzing verbal responses, tone of voice and even facial expressions, such as the tool "HireVue". This tool allows employers to retrieve recorded interview responses from job applicants and then rate these responses against the interview responses of current, successful employees. More specifically, HireVue's tool analyzes videos using machine learning and extracts signals such as facial expressions, eye contact, vocal signs of enthusiasm, word choice, word complexity, topics discussed and word groupings. According to Miranda Bogan and Aaron Rieke (2018) the use of tools such as "HireVue" raises ethical questions. They think speech recognition software may perform poorly, especially for people with regional and non-native accents. Additionally, voice recognition still has significant race and gender biases (Palmiter Bajorek, 2019). Furthermore, facial analysis systems may have difficulty recognising faces of women with darker skin tones. In addition, some respondents might be rewarded for irrelevant or unfair factors, such as exaggerated facial expressions, and penalized for visible disabilities or speech impediments. On the other hand, the use of this type of biometric data might have no legal basis if the data is used to predict success in the workplace, to make or inform hiring decisions.

There are hiring tools that currently aim to predict whether applicants might violate workplace policies or to assess what mix of salary and other benefits should be offered. The fear is that such tools could widen the pay gap for women and non-white workers (Bogan and Rieke, 2018). This is because data obtained by Human Resources departments tends to contain a lot of information on a future worker's socioeconomic and racial status, which could be reflected in predictions of salary requirements. Furthermore, offering employers highly specific insights into an applicant's salary expectations increases the information asymmetry between employer and applicant at a critical moment in the negotiation.

Another overall ethical standard that has been agreed on in order to avoid bias in AI is fairness. In this context, fairness means a set of procedures aiming to avoid bias so as to ensure outcomes that respect ethical standards such as acknowledgement of human agency, privacy and data governance, individual, social and environmental wellbeing, transparency and accountability, and oversight (Xenidis and Senden, 2020).

Many authors agree that AI has to be safe. By way of example, the center for AI safety at Stanford has the intention to lead this important aspect of AI Ethics by developing "rigorous techniques for building safe and trustworthy AI systems and establishing confidence in their behaviour and robustness, thereby facilitating their successful adoption in society" (Barrett et al). Safety in AI plays especially a role in robots, autonomous vehicle and security issues such as the above-mentioned LAWS (Coeckelbergh, 2020).

According to AI HLEG (2019) "the data, system and AI business models should be transparent. Traceability mechanisms can help achieving this. Moreover, AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned. Humans need to be aware that they are interacting with an AI system, and must be informed of the system's capabilities and limitations." For example, humans should know if they speak to a conversational AI such as AI voice assistants. The transparency is important on the one hand, because the human being might be used for the machine learning of the AI voice assistant and there arise privacy and data protection issues. On the other hand, humans have a right to understand the decision taken by the emerging AI technologies (see Coeckelbergh, pages 116-123). This is not always the case and leads to the so-called black box problem. This means that there are some decisions of AI based decisions that cannot be explained. This can lead to further procedural difficulties in allocating responsibilities within a fragmented chain of actors and complex human-machine interactions and attributing liability across multiple legal regimes in complex and composite AI systems.

Another important ethics issue related to AI is accountability. According to AI HLEG (2019) "(m) echanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data and design processes plays a key role therein, especially in critical applications. Moreover, adequate an accessible redress should be ensured."

AI RELATED HUMAN RIGHTS

As mentioned above, ignoring AI Ethics can lead to human rights violation. According to Articles 7 and 8.1 of the Charter of Fundamental Rights of the EU, "(e)veryone has the right to respect for his or her private and family life, home and communications" and "(e)veryone has the right to the protection of personal data concerning him or her. These rights to respect for private life and the protection of personal data are one of the main subjects in fundamental rights discussions around the use of AI.

For the machine learning of emerging AI technologies, developers need a lot of data to make the AI models reliable and effective. As more humans participate in the feeding of machine learning the better for the AI technologies. The EU Agency for Fundamental Rights (2020) interviewed AI developers and asked them about their use with personal data. In the interviews, "respondents were not always entirely clear about their use of personal data." In many cases, respondents explained that data protection was not relevant in their activities because they used non-personal data or anonymised data. One parastatal company in the environmental management sector explained that it uses aggregated water consumption data for machine learning-based water consumption forecasts. According to the interview partners such data cannot be collected at the individual level. Other interviewees explained that they do not use data collected from individuals as personal data. For example, the inspection of restaurants with their collection of data from online sources does not use personal data. On the other hand, these interviewees admitted that they were cautious about online data collection because, even if publicly available, it could contain personal data such as usernames.

Another problem might be more related to Article 1 of the UDHR "(a)ll human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood." For many years, users of applications such as Google maps, Facebook, WhatsApp or other social media have been contributing to the feeding of the machine learning industry. As it is explained in the new Netflix documentary "The Social Dilemma", not personal data itself is the new gold, however it is a gold rush for personal data to build more and more accurate profiles around us. Shoshana Zuboff invented the term "surveillance capitalism" in 2014 to describe the methods technology firms use to claim our private experiences and turn them into their products. In her book "The Age of Surveillance Capitalism" she describes instrumentarianism which is a new kind of unaccountable power that surveillance capitalism has created. "This means the instrumentalization of behavior for the purposes of modification, prediction, monetization and control that threatens to challenge some of the functions of the state and usurp the sovereignty of the people. Instrumentarianism can determine the ends, because it can manipulate the means" (Thornill 2019). This means that every user of AI based technologies should be aware of the fact that also can be part of the instrumentarianism, and then everybody has to decide if she or he wants to be part of this experiment by using or not some of these tools.

Equality and discrimination are crucial topics when it comes to the use of AI, because the very purpose of machine learning algorithms is to categorize, classify and separate.

According to articles 21.1 and 23 of the EU Charter of Fundamental Rights "(a)ny discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited" and "equality between women and men must be ensured in all areas, including employment, work and pay".

There are many examples of algorithmic discrimination. For example, a Google search on "professional" hair show mostly pictures of white women while a search for "unprofessional" hair displays predominantly pictures of black women. Similarly, facial recognition applications perform much worse at recognising black women's faces than white men.

In 2019, a first instance decision of the Divisional Court of Cardiff dismissed a claim concerning the lawfulness of the South Wales Police's use of the "AFR Locate" face recognition system. The Court of Appeal overturned that decision, because it found that the facial recognition programme used by the police was unlawful. The Court of Appeal ruled that "too much discretion is currently left to individual police officers". It added that "[i]t is not clear who can be placed on the watch list, nor is it clear that there are any criteria for determining where [the technology] can be deployed". (UK, Court of Appeal, R (Bridges) v. CC South Wales, [2020] EWCA Civ 1058, 11 August 2020; Ars Technica, 'Police use of facial recognition violates human rights, UK court rules', 11 August 2020.) Furthermore, the judges held that the police did not sufficiently investigate if the AI based technology exhibited race or gender bias. In Europe, this is the first ruling specifically on AI and algorithm discrimination. For law enforcement agencies, it significantly narrows the scope of what is permissible and what must be done to fully respect human rights.

Meredith Broussard (2019) has already been warning in 2014 "Why Poor Schools Can't Win at Standardized Tests". She uses as example a Pennsylvania System of School Assessment. (PSSA) test from 2009 for students aged 8-9 and the question to write down an even number with three digits and to explain how they arrived at their answer. The following answer is correct: "932 is an even number because all you have to do is look in the ones place. If you can divide that number equally, it is even. An even number in the ones place makes the whole number". If a student answered: "200, cause it sis an even number and it has digits" the answer would only be partially correct, as "the third-grade student lacked the specific conceptual underpinnings to explain why it is correct" (Broussard, 2019). Unfortunately, not all students have access or can afford to buy the books which provide this very specific knowledge on this very specific answer.

The origins and causes of algorithmic discrimination reflect existing discrimination in our offline, real world where humans discriminate against each other. It starts with our own human stereotypes that have led to discrimination in the past (such as men are strong and women are weak, or assumptions about racial stereotypes). The consequences are structural inequalities. Stereotypes and biased conduct enter—consciously or unconsciously— into the design of an algorithm. This leads to the generation of biased data, such as in the cases explained before.

If societies want to avoid repeating the same patterns of bias and discrimination that we witness in the 'physical world', algorithmic discrimination needs to be addressed. Otherwise, humanity risks creating a digital world that replicates structural inequalities provides.

FUTURE TRENDS AND ISSUES AI ETHICS AND HUMAN RIGHTS

The members of the AI HLEG (2019) outline their concerns related to the future of AI. Firstly, the experts are concerned that AI is enabling increasingly efficient identification of individuals by both public and private entities. They list a number of scalable AI identification technologies, such as facial recognition and other involuntary identification methods using biometrics (i.e., lie detection, micro-expression personality assessment and automatic voice recognition). The AI HLEG (2019) is well aware that identification of individuals is sometimes the most desirable outcome consistent with ethical principles (for example, in detecting fraud, money laundering or terrorist financing). However, this is countered by the fact that automatic identification can have unexpected effects on many psychological and socio-cultural levels, some of which humanity is currently unable to predict. For this reason, a proportionate use of control techniques in AI is necessary to preserve the autonomy of citizens. It must also be clearly clarified whether, when and how AI can be used for the automated identification of persons. Furthermore, a distinction should be made between identifying a person versus tracking and tracing a person, and between targeted surveillance and mass surveillance.

The second aspect was already mentioned above, namely that people should always know whether they are interacting directly with another human or a machine. The experts of the AI HLEG (2019) see AI practitioners as responsible for ensuring that this is reliably achieved. For this reason, AI practitioners should ensure that humans are made aware of - or are able to request and confirm - the fact that they are interacting with an AI system (e.g., by issuing clear and transparent disclaimers). The experts are aware that there are borderline cases that complicate the matter (e.g., an AI-filtered voice spoken by a human). However, it should be borne in mind that the confusion between human and machine can have a variety of consequences, such as binding, influencing or diminishing the value of being human. For these reasons, the experts believe that the development of human-like robots should be subject to careful ethical evaluation.

In third place, the experts of the AI HLEG (2019) see societies as having a duty to protect the freedom and autonomy of all citizens. They assume that any form of citizen assessment can lead to the loss of this autonomy and endanger the principle of non-discrimination. Above all, the screening of people, i.e. scoring, should only be used if there is a clear and legal justification. Furthermore, measures must be proportionate and fair. When public authorities or private actors engage in normative citizen scoring (general assessment of "moral personality" or "ethical integrity") in all aspects and on a large scale, human rights violations may occur; especially when scoring is not used in accordance with fundamental rights, and when it is used disproportionately and without a delineated and communicated legitimate purpose.

The experts explain that citizen scoring - to a greater or lesser extent - is already frequently used in purely descriptive and domain-specific scoring (e.g., in the example mentioned above with the American PSSAs, school systems, e-learning and driver's licences). Again, citizens should be provided with a fully transparent process. They should be given information about the process, purpose and methodology of the assessment. However, the experts rightly point out that transparency cannot prevent non-discrimination or ensure fairness, and it is not the panacea to the problem of assessment. For even if citizens know that they are being evaluated, they cannot always understand on what basis this is done. Therefore, it should be possible for citizens to opt out of the scoring mechanism without suffering any disadvantages. Otherwise, the experts think that there should be mechanisms to challenge and correct the scores. This last aspect is of high importance if there is an asymmetry of power between the parties.

At all times, such opt-out options should be ensured in the design of the technology in order to ensure that fundamental rights are respected. This is necessary in a democratic society.

The fifth aspect that worries the experts of the AI HLEG (2019) is the development of the above-mentioned LAWS and they urge for regulation in this regard.

Lastly, experts still see the development of AI as domain-specific, requiring well-trained human scientists and engineers to precisely specify its goals. However, this could change in the future. For this reason, the experts suggest that we should be prepared for problems and challenges that are as yet unthinkable, in order to be able to react to possible unknown unknowns and "black swans". This also includes a regular assessment of these issues.

TEACHING AI ETHICS

Artificial intelligence1 (AI) growing at a really fast pace. We should expect to see significant modifications in our society as AI systems emerge as embedded in many components of our lives. AI will have a central role in daily in the near future. However, its autonomous decision making feature and complex nature of AI powered systems pose considerable regulatory and legal challenges. The Oxford Dictionary ("Ethics," 2021) defines ethics as a set of moral principles that govern a person's behavior or the conduct of an activity. Professional ethics refers to the moral issues related to the specialist knowledge that professionals gain, and how the use of this knowledge should be governed when providing a service to the public (Chadwick, 1998).

Ethical considerations of AI applications attract great interest in the society. International institutions and states are working on issues related to the ethical use of AI technologies and tools. Hence, European Union published a report titled "Recommendations to EU Commission on Civil Law Rules on Robotics" and the USA released a National Artificial Intelligence Research and Development (NAIRD) Strategic Plan (Furey & Martin, 2019).

With AI powered machines we are creating a new kind of mind. AI is getting more pervasive in our Daily life. There is a need to consider what the birth of the new mind might mean in our Daily interactions and how it may shape us and how we might shape it. Can we give to an artificial mind a sense of human morality and ethical values? It is almost impossible to predict the answer of the questions: can AI powered advance machines take over their own evaluation with a machine mind and go beyond our minds and decides that human is a treat for them.

Nobody knows the answer of this question at this point. Worrying about the future with AI powered machines will not help us to prepare our students to the future. Instead we should prepare them how to tackle with problems that might be caused by AI powered machines in our daily life.

We will see more computers and machines equipped with AI in the near future. Companies design new machines that are capable of making their own decisions such as autonomous cars, robots and drones. In practice, these autonomous systems will need moral guiding programs for certain conditions to take an appropriate moral decision. However, in real life we face moral dilemmas to make the best moral decision. The decision we make has also consequences. It is also true for the autonomous systems. For instance, a self-driving car should be instructed to change its direction to avoid hitting an object in order to protect its passengers even if such decision may lead to hit a vehicle in another lane. This is an example of the ethical dilemma that we will face in the near future. Therefore, schools should address these ethical issues and prepare them for tomorrows' world.

AI Ethics is a new field of study that requires a new ethical thinking. There are many questions that arise when we think about ethics of AI. Such as, how do we prevent learning algorithms from obtaining morally objectionable biases? Should autonomous AI powered war machines be used to kill in battle? How should AI powered technologies be inserted in our social relations? Is it acceptable to fall in love with an AI system? What kind of ethical rules should AI have powered autonomous car use? Can AI systems be affected moral abuses? And if so, of what kinds? Can AI systems be moral agents? If so, how should we hold them accountable? How should we live with and tolerate minds that are incongruous to our own?

AI ethics teaching cover these questions and related topic that students should be able to exhibit knowledge of philosophical issues involved in ethics of AI, demonstrate familiarity with relevant examples of AI powered technologies, exhibit ability to express arguments clearly and concisely related to ethical issues of AI and gain skills in research, analysis and argumentation related to AI ethics and moral issues of AI.

Ethical concerns Stahl and Coeckelbergh (2016) identify ethical and social issues of concerns related to AI. These issues are identified as:

- Trust: Shall we trust giving care by AI powered robots?
- Data protection and privacy: which data are collected? How the data stored? Who has access to the data? Who owns the data?
- Safety and avoidance of harm: AI powered machines should not harm people and be safe to work with.
- Deception: If IA powered machines and technologies are used as "social companions", are these roles can be considered as deception?
- Responsibility: When the AI powered machines take responsibilities of human tasks, who will take responsibility of the ethical issues caused by AI powered machines?
- Moral agency: AI powered machines and technologies do not seem to have the capacity of moral reasoning or, more generally, of dealing with ethically problematic situations. AI power machines does not seem to have the capacity of moral agency when and human interaction caused a moral problem. AI powered technologies have no capacity to understand and think moral and ethical norms of the society. There are two criteria proposed for moral status. There criteria are sentience and sapience (Bostrom & Yudkowsky, 2014). Sentience refers to the capacity to feel suffer and pain. Sapience refers to being a reason-responsive and self-awareness. Human beings have both sentience and sapience capacities. However, AI powered machine have not reached these capacities to think and act based on ethical norms of the society.
- Replacement and its implications for labour: Are AI powered technologies developed to solve problems in real life or are they developed to save money by replacing human?

Employing AI powered technologies and tools has become increasingly widespread in daily life. Additionally, there is a deep concern about the societal impact of use of AI. Therefore, there is an urgent need to prepare students ethical and moral implications of AI powered machine and tools. Although, AI technologies make our daily life more efficient, they can also result in problematic results such as misinformation, wrong decisions, and biases. It is vital for future generations to be aware of the pitfalls of AI powered technologies in order to make sure that all current and future AI technologies serves us for good.

Today's students are tomorrows workforce. They have to understand how to work with AI powered technologies. It is educational institutions' duty to empower and teach them in learning ethics and legal issues of AI powered technologies. Furthermore, we should also raise students' awareness about ethical and legal issues embedded within the way AI powered technologies. Students should learn that AI powered technologies are developed by human being. They should aware of the fact that AI powered technologies are not just robots they are technologies that run on AI algorithms coded by humans. This means that those AI powered machines are likely to make the mistakes and to have biases as human beings.

Teaching AI ethics and moral may require basic information about what is AI and how it works. Therefore, teacher should start from explaining basics of AI before start teaching AI ethics and moral issues related with AI technologies. Furthermore, basic vocabulary of AI should be given to students before teaching AI related issues. Case studies and sample scenarios are good tools to use teaching AI ethics. Teacher may ask questions to students related to scenarios with ethical issues of AI to open discussions. Discussions allows students to think and analyse possible AI related ethical issues and prepare answers based on their ethical and cultural background.

DISCUSSION/CONCLUSION

Over the past decade, AI powered technologies developed exponentially and become our daily life tool across entertainment, autonomous cars, ecommerce, personal assistants, and advertisement. However, there is a danger awaiting ahead of our times. The misuse of AI powered technologies, moral and ethical issues of AI. The potential misuse of AI and ethical issues that may cause by AI technologies become cause of concern. There is an increasing demand to establish ethical standards and regulations of AI use. In addition to this urgent demand, there is also a need to teach ethical and moral issues which may be caused by AI powered technologies to the future generations. It is important to decorate future generations with AI ethics competencies to critically reflect on their learnings in their future jobs.

REFERENCES

Barrett, C., Dill, D. L., Kochenderfer, M. J., & Sadigh, D. (n.d.). *Stanford Center for AI Safety. White Paper.* Available in http://aisafety.stanford.edu/

Bogen & Rieke. (2018, Dec.). Help Wanted. An Examination of Hiring Algorithms, Equality and Bias. *Upturn*, 36-38.

Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *The Cambridge Handbook of Artificial Intelligence* (pp. 316–334). Cambridge University Press. doi:10.1017/CBO9781139046855.020

Broussard, M. (2019). *Artificial unintelligence: How computers misunderstand the world*.

Chadwick, R. (1998). Professional Ethics. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*. Routledge.

Coeckelbergh, M. (2020). *AI Ethics. MIT Press*.

Coeckelbergh, M., & Stahl, B. C. (2016). Ethics of Healthcare Robotics: Towards Responsible Research and Innovation. *Robotics and Autonomous Systems*, *86*, 152–161. doi:10.1016/j.robot.2016.08.018

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. Available at https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Ethics. (2021). *Oxford English Dictionary*. Retrieved from http://oxforddictionaries.com/definition/english/ethics

EU Agency for Fundamental Rights. (2020). *Getting the Future Right*. Artificial Intelligence and Fundamental Rights.

Furey, H., & Martin, F. (2019). AI Education Matters: A Modular Approach to AI Ethics Education. *AI Matters: A quarterly newsletter of the ACM Special Interest Group in Artificial Intelligence*.

High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. Available in https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Kant, I. (1785). Grundlegung zur Metaphysik der Sitten, 1785. In *Immanuel Kant, Gesamtausgabe in zehn Bänden, Leipzig 1838, Modes und Baumann* (p. 23). Academic Press.

Mayor, A. (2018). Gods and Robots. In Myths, Machines and Ancient Dreams of Technology. Princeton University Press. doi:10.2307/j.ctvc779xn

O'Neil, C. (2016). Weapons of Math Destruction. Penguin Random House LLC.

Palmiter Bajorek, J. (2019). Voice Recognition still has significant race and gender biases. *Harvard Business Review*. Available at https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases

Pieper, A. (1991). *Einführung in die Ethik*. UTB.

Thornill, J. (2019). Should we think of Big Tech as Big Brother? *Financial Times*. Available at https://www.ft.com/content/43980f9c-0f5b-11e9-a3aa-118c761d2745

United Nations Human Rights Office of the High Commissioner. (n.d.). https://www.ohchr.org/EN/Issues/Pages/WhatareHumanRights.aspx

Vieth, A. (2006). *Einführung in die Angewandte Ethik*. WGB.

West, D. M., & Allen, J. R. (2020). *Turning Point. Policymaking in the Era of Artificial Intelligence*. The Brookings Institution.

Xenidis, R., & Senden, L. (2020). EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination. In General Principles of EU law and the EU Digital Order. Kluwer Law International.

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. Hachette Book Group.

KEY TERMS AND DEFINITIONS

Algorithm: Algorithm refers to set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

Artificial Intelligence (AI): AI refers to the simulation of human intelligence in machines that are programmed to think and act like humans.

Coding: Coding refers to the language used by computers to understand our commands and, therefore, process our requests.

Ethics: Ethics refers to moral principles that govern a person's behaviour or the conducting of an activity.

Machine Learning (ML): ML an application of AI it gives devices the ability to learn from their experiences and improve their self without doing any coding.

Section 3
Digital Ethics in Health Care

Chapter 6
Digital Equity:
Responding to the Reality of the Digital Divide

Patrick Flanagan
St. John's University, USA

ABSTRACT

This chapter discusses digital equity through the lens of the digital divide. While the digital divide is as old as information communication technology itself (ICT), the COVID-19 health crisis renewed a strident interest in exposing the significant gap that still exists after close to 30 years. The digital divide then is first contextualized within the coronavirus pandemic to illustrate how inequities came further to the forefront of people's agenda. It then moves to discuss the digital divide defining the complex term and offering critical data to illustrate the areas of the world most impacted by this unfortunate reality. Different organizations and groups have made significant moves to narrow the digital gap. These strategies are discussed next. None of these groups will be fully successful if, as will be argued, they are not concerned with digital equity. Finally, the chapter makes some critical observations on future challenges facing ICT vis-à-vis the digital divide.

INTRODUCTION

In the early months of 2020, the global community faced an unprecedented historical reality not experienced on such a scale since the AIDS epidemic unfolded in 1981. COVID-19 was unlike more recent significant world health challenges. Previous pandemics paralyzed the world community in unimaginable ways, but this latest one would offer a bit of a breather. With the advances in information communication technology (ICT) over the past three decades, life could and would continue for some in what analysts would identify as "the new normal." This forecasting was in no way to diminish the devastating impact the coronavirus had had on the disruption of life and the subsequent destruction it caused in so many already vulnerable parts of the world. Instead, it acknowledged that while out of an abundance of caution, governments civilly legislated lockdowns, distancing, and the wearing of masks and citizens personally

DOI: 10.4018/978-1-7998-8467-5.ch006

obliged, peoples' access to their computers and networks made life a bit more bearable for some. Unfortunately, not all in the global village enjoyed the refuge of pivoting to information technology for their work, school, and social life. While the move to online work was a novel transition, at least at first, the raw reality was and continues to be a lack of access to the Internet for several citizens in the global village.

COVID-19 ravaged lives, particularly for those with comorbidities, and the deadly virus profoundly affected those who loved them. Without question, though, the very threat of contracting the pathogen challenged an even more extensive section of the global population. It has been such a deflating anxious experience for so many. People had more time on their hands with the luxury of not having to do as much commuting, and it allowed them to critically assess injustices not so evident before the pandemic. So much of the soft underbelly of global economic, political, biomedical, religious, and social systems were laid bare. Some people took to the streets in protest; others used social media and Websites to highlight the inequities they perceived in historical realities. For almost half of the world's population, the latter move was not an option. They remain disconnected from the Internet. Although others have laid claim to the phrase, it was Sir Winston Churchill, former Prime Minister of the United Kingdom, who in the mid-1940s, as the end of World War II was approaching, "never let a good crisis go to waste." Citizens of the world have capitalized on this during this uncertain and debilitating historical period. Activists, old and new, have called leadership from all realms of life to accountability for their leadership. For example, in sections of Europe and the United States, crowds of people took to the streets to protest what they understood to be an inordinate amount of violence towards and disregard for people of color, especially law enforcement personnel. In America, particularly the "land of the free and the home of the brave," the significant pause in life due to the coronavirus has proffered opportunities to reassess the nation's commitment to "justice for *all*."

The lack of digital justice or, more poignantly, the digital divide increasingly became a concern as the COVID-19 virus penetrated more and more parts of the world. As commerce, education, worship, and social life pivoted to virtual experiences, workers, students, and consumers needed equipment to maximize their lives. Corporations and schools outfitted their staff and students with computers. Businesses and educational institutions proudly provided end-users with needed devices did not anticipate the number of users who did not have network access.

DEFINING THE DIGITAL DIVIDE

The digital divide is a term that describes the gap between those who have access to information communication technology (ICT) and those who have limited or no access. This distinction, however, between the "haves" and the "have nots" can be too fundamental a delineation (Compaine, 2001; Hawkins, 2006; Selwyn, 2004; Warschauer, 2002). What is "had" and "not had" is much more comprehensive, involving available physical equipment, utility resources (for instance, electricity), and technological skills. While the "have nots" can be those who do not have adequate access to information communication technology, the "haves" can include those who have a computer, but with no or limited connection to the Internet, with a rather dated dialup and not a broadband connection, or those who connect through a mobile phone. ICT has significantly transformed political, social, and economic engagement in connected parts of the global village. Without effective widespread access to ICT, the digital divide further alienates citizens within and among countries of the world and amplifies divides already established ethnic, gender, income, and geographic inequalities. Both government agencies and scholars have studied the digital divide carefully and have suggested creative ways to ensure access to equipment, education, and viable signal connections to maximize fuller participation in this dynamic global ICT phenomenon.

A review of literature early on in the rollout of the WWW reveals attentiveness to more than just lack of access to the rich technological resources some enjoyed. In their assessment of the digital divide, scholars highlight that the chasm is much more complex than its original sense involving widespread inequalities on various political, economic, educational, demographic, ability, and gender levels (Alampay, 2006; Barzilai-Nahon, 2006; Colle and Roman, 2001; Dagron, 2001; DiMaggio, Hargittai, and C & S, 2004; Fink and Kenny, 2003; Norris, 2001; Parkinson, 2005; Potter, 2006; Simpson et al, 2004; and Warschauer, 2003). While admitting, for example, the excitement of the Internet's impact for optimizing networking in the global village, Norris (2001) raised some critical questions as to whether or not the Internet would evolve into a democratic, participatory medium offering equal advantages for engagement or would it only reinforce dominance and inequality. Beyond a binary construction of the digital divide rendering it more complex, Norris describes three divides that called for a response: the global divide that focused on access; the social divide that alienated people; and the democratic divide that illustrated the use or lack of use of the Internet for political purposes. Van Dijk and Hacker (2003) identify psychological, material, skill, and usage factors influencing this access. Hilbert (2004) focuses on the gender divide while Preiger and Hu (2006) study the racial divide, both further specifications of the digital divide. Kularski and Moller (2012) further specify the digital divide focusing on the technological skill gap. Kularski and Moller note that the challenge involves more than supplying ICT equipment and ensuring access points to digitally excluded people. Users need to be trained on how to use technology optimally for their needs.

Building upon a 2019 Pew Research Center study, Perrin, Lai, and Widmar (2021), discuss more fully the lack of broadband services in rural areas of the United States amplified during the coronavirus pandemic. In this age of restricted movements, telehealth became an increasingly popular meeting point for medical personnel and patients. However, as Ramsetty and Adams (2020) pointed out, the lack of online connections or unstable ones, at best, precluded some from responding promptly to pressing health concerns. Those successful connections only further revealed the complexity of the digital divide. It is not merely about access but also concerns access quality, age, income, race, and language.

MAPPING THE DIGITAL DIVIDE

Since 1993, the website Internet World Stats (www.internetworldstats.com) has been tracking many variables associated with the development of the Internet and citizens' engagement. In 1993, out of the 5,578,865,110 people in the world, there were only 14,161,570 Internet users. Thus, a mere 0.3% of the global village enjoyed a connection. In the first quarter of 2021, 64.2% of the world's 7,875,765,584 people, or to be more exact, 5,053,911,722, are connected to the Internet. While a considerable growth, the reality remains that 35.8% of the world's citizenry remains untethered to the technology that so many enjoy.

One of the Internet World Stats resources on its site, "The Internet Big Picture," offers a window into where the marked increases are happening in the world. While all continents demonstrated growth, the data reveals that Africa, Asia, and Oceania/Australia still have a steep climb to make the connections to the Internet that North America and Europe enjoy.

Table 1. World internet usage and population statistics, 2021 Year-Q1 estimates

WORLD INTERNET USAGE AND POPULATION STATISTICS						
2021 Year-Q1 Estimates						
World Regions	**Population (2021 Est.)**	**Population % of World**	**Internet Users 31 Dec 2020**	**Penetration Rate (% Pop.)**	**Growth 2000-2021**	**Internet World %**
Africa	1,373,486,514	17.4%	590,296,163	43.0%	12,975%	11.7%
Asia	4,327,333,821	54.9%	2,707,088,121	62.6%	2,268%	53.6%
Europe	835,817,917	10.6%	728,321,919	87.1%	593%	14.4%
Latin America / Caribbean	659,743,522	8.4%	477,869,138	72.4%	2,544%	9.4%
Middle East	265,587,661	3.4%	188,132,198	70.8%	5,627%	3.7%
North America	370,322,393	4.7%	332,919,495	89.9%	208%	6.6%
Oceania / Australia	43,473,756	0.6%	29,284,688	67.4%	284%	0.6%
WORLD TOTAL	7,875,765,584	100.0%	5,053,911,722	64.2%	1,300%	100.0%

The websites Statista (www.statista.com) and Internet Live Stats (www.internetworldstats.com) offer a deeper dive into usage demographics. 2.7 billion of the 5.05 billion Internet users lived in 10 countries. The percentage of internet users given the populations of countries in order are: China (69.3%), India (57.7%), United States (95.5%), Indonesia (79.3%), Brazil (76.4%), Nigeria (69.5%), Japan (91.5%), Russia (85.1%), Mexico (69%), and Iran (95.5%). The other half of the geographic picture from the bottom up are the following countries: North Korea (0%), Eritrea (6.9%), South Sudan (8%), Comoros (8.5%), Central African Republic (11.1%), Somalia (12.4%), Burundi (13.3%), Niger (13.6%), Kiribati (14.6%), and Liberia (14.9%).

This data vis-à-vis least connected countries will come as no surprise and are rooted in a lack of available resources and the dominant political narrative. Connections translate as business profit as well the opportunity for the advancement of democracy. For these reasons and others, the digital fissure so many of the world's people still experience remains of interest to global leadership in both these fields. Rightsizing the inequalities in technology can increase commerce and build a global community rich with solidarity, particularly concerning human rights. Suppose these technological deficits persist and are not resolved with some reasonable alacrity. In that case, only further digital inequalities would continue to make it difficult for least connected countries to compete, participate, resource, and communicate the information superhighway.

NARROWING THE DIGITAL DIVIDE

The data reveals that over the past almost three decades now, the global community is more connected. Through innovation on the part of both for-profit and not-for-profit organizations, there are more people online. This praise of corporations and charities does not deny the priority that governments have made to close the digital divide. Evident progress has been made to get people online. No enterprising solution is a solitary accomplishment. There are three recent creative solutions that evidence the partnerships that are made to close the digital gap and get people connected.

The most prominent global solution can be found in the United Nations' 2030 Agenda for Sustainable Development (Carpentier & Braun, 2020). In 2015, all member states committed themselves to the work this document proposed. In the Agenda's 17th goal, the United Nations made technology a priority in this "blueprint for peace and prosperity for people and the planet, now and into the future." United Nations members believed that poverty would continue to exist and possibly exacerbate if people continue to experience inaccessibility to the Internet. Admittedly, it can be understood as an unrealistic goal given where the global community has been vis-à-vis technology. Still, with the United Nations naming as one of its goals, the lengthy journey towards ending the digital divide may be in reach (Sparviero & Ragnedda, 2021).

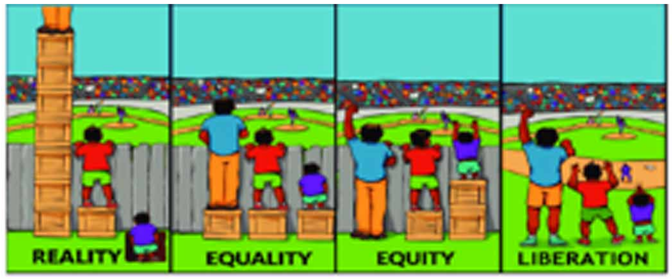
Close the Gap is a more precise example with a history since 2003 that evidences sustainable impact. Its mission is clear. As an "international social enterprise," Close the Gap offers "high-quality, pre-owned computers donated…to educational, medical, and social projects in developing and emerging countries. A review of their website (www.close-the-gap.org/) offers visitors an appreciation of the contributive reach Close the Gap had in its almost two decades. They received 1,007,000 computers, supported 6,280 projects, and 3,220,000 people have benefited from their efforts. This charity distinguishes itself not only in its achievements but also in its efforts to reduce e-waste. Perhaps, most notable though, is the fact that its efforts are "demand-driven and impact-oriented." For all its rightly celebrated accomplishments, any knowledgeable observer has to recognize there are problems with sustainability, particularly in the global South.

Previously, we noted that the digital divide runs deeper than the mere macro level of who has the access and who does not. Grow with Google, an initiative of Alphabet Google has successfully transformed the deeper levels of the digital divide since its foundation in 2017. People may have access to broadband options, but knowing how to maximize engagement and accessing the ideal resources is not always known by users. This corporate commitment continues to provide free assistance with expert training, tool selection, and skilly talent. The statistical impact is significant. Five million Americans have advanced their skills, careers, and businesses because of this training. In addition, over 250,000 people have enrolled in Google's IT Support Professional Certificate program. Grow with Google has translated into $416B in economic development for users optimizing Google products. Through this Grow with Google program, participants become hungry to be life-long learners and regularly seek new growth opportunities. This Alphabet Google program has proven critical for members of historically marginalized communities.

EQUITY

After the historical death of George Floyd amid a time of heightened sensitivity because of the pandemic, the global community galvanized itself to fight for systemic racial justice. In addition to the protests and marches throughout the world, businesses, schools, and churches started or renewed efforts to ensure diversity, equity, and inclusion. While many people might have had an appreciation of the end words, equity, for some though, translated into equality. How far from the precise definition of the word.

Figure 1. The difference between the terms equality, equity, and liberation, illustrated
© Interaction Institute for Social Change | Artist: Angus Maguire



"Equity" is a word recent to our lexicon and used more frequently in contemporary conversations. A Google evaluation of the word's use indicated that it was frequently used in the middle of 2020, most notably in conjunction with the strident efforts of activists to bring attention to the inequitable treatment of Black people. Equity is a clarion call to provide access to those who have historically been sidelined or blocked. The artist, Angus Maguire from the Interaction Institute for Social Change, illustrated it best in artwork entitled "The Difference Between the Terms Equality, Equity, and Liberation (2020)."

In Maguire's depiction, he shows the progression from reality to liberation. This is apropos for describing the reality of the digital divide. At present, we recognize that some have more access to ICT than others. And others have nothing. While those "with" are speaking about 5G speeds, those who are reaching or without are struggling to enjoy ICT.

It is not a matter of making the platform equal. Such a move only exacerbates the problem of patronizing those without and keeping them in "the dark." Equity offers the least, the last, and the lost the necessary boosters to assist them in enjoying the digital connections others have enjoyed. Ideally, for true liberation, no blockages should exist, but equity is a critical step.

While the pandemic shattered people's lives, it also moved the world community to an essential reevaluation of many ways it was relating. The rawness of reality that proceeded the onslaught of Covid-19 offered time for significant reflection and strategizing to resolve injustices. The prospect of building a society and engaging ICT based on equity offers a much more comprehensive and profound tactic for resolving the digital divide.

THE INCENTIVE FOR NARROWING THE DIGITAL DIVIDE

Belarus, Burma, China, Cuba, Egypt, North Korea are the top five countries that limit ICT access. For those who have ease of use of ICT, it may be perplexing. Essentially, ICT truly is the World Wide Web, introducing users to material both familiar and foreign. However, to these five and a host of other countries, information on the Internet does not gel with the narrative these governments wish to preserve. They seem to be the reincarnation of the legendary Ned Ludd, who believed technological advancement was a threat to society.

Technology is transformative. Amid the Covid-19 pandemic, political movements gained force and garnered sympathizers because of their presence on ICT. The move to address systemic racism, economic inequality, government transparency, and healthcare availability was swift using the available social media tools on ICT. Access to the Internet can imagine “a world of carnival, community, and contention in and through cyberspace and how in this process they have transformed personhood, society, and politics” as “a response to the grievances, injustices, and anxieties caused by the structural transformation of Chinese society” (Yang, 2009). Without access to ICT, nothing changes, and inequities have the capacity to persist.

ICT is not the salvo to the myriad of concerns facing the global community. However, ICT does engender solidarity and subsidiarity. ICT permits users to imagine a different world and take steps, sometimes in partnership with those with more resources, to resolve depraved conditions. At a minimum, though, it all begins with a decent device that can access ICT in a sustainable way. ICT may not be the universal remedy, but it can be a substrate for resolving economic, political, and social challenges.

FUTURE RESEARCH DIRECTIONS

Resolving the digital divide requires a multi-dimensioned response and will continue to entail a great deal of strategic analysis and priority planning. Further research to study carefully many of the critical concerns raised in this chapter will be essential to sustain any momentum to close the digital gap. While the penetration of mobile devices throughout the global village has been hailed as a significant advancement in bridging the digital divide, mobile telephony is only a step. A person using the mobile phone for Internet access in its present form cannot fully engage the broader resources on the Web; a larger Internet-connected computer or tablet is ideal. In addition to securing physical provisions for technologically handicapped nations, access to reliable and economically reasonable Internet signals is also as important.

Another area of future research must involve ongoing training and development. Internet-connected computers cannot merely be given to digitally challenged nations with the expectation that they will be able to maximize the use of technology. While the technological learning curve can be thin in younger populations, who are more willing to take risks and learn through experimenting with technology, strategic efforts to ensure training and subsequent updating of the wide range of computer users will be critical. As advancements in computer technology become available, it would also be necessary to regularly do work to supply the more marginalized and vulnerable communities with contemporary machinery, stronger Internet signals, and updated software. In the latter case, the Open Source Movement offers marginalized peoples opportunities to use standard software programs and, at the same time, participate in tweaking the software (Bergquist, Ljungberg, & Rolandsson, 2011).

A final area that researchers need to study more deeply focuses on the motivations of leaders and nations to restrict access to ICT. The plethora of information and connections offered through ICT can maximize global engagement and dynamic innovation. At the same time, some leaders of nation-states are threatened by the resultant thoughts and actions after interacting with the massive capabilities the Internet offers. Research into how cultures have advanced through their engagement with ICT and how Luddite countries might move beyond its technological paralysis would be helpful to both individual citizens and the global community.

CONCLUSION

The Covid-19 has ravaged communities in ways unimaginable. In such a time of stay-at-home regulations, citizens have had much time to take stock of what they have and do not have. To stay connected through the impressive power of technology, whether it was medically, educationally, professionally, or socially, some have recognized the digital impoverishment that so many experienced. ICT has not been the invaluable construct for all its celebrated advantages, so many have claimed it to be. This reality is for no other reason is that there continues to be a significant portion of the population that remains digitally disconnected. While those with digital access, there may be other problems such as skillsets, utility access, and sustainability of equipment, to name a few challenges.

Like efforts during the pandemic, government and health officials made efforts to "flatten the curve." So, too, similarly must be done vis-à-vis the digital divide. While steady efforts have been made to crush the abscess since the introduction of ICT to the public, the global community may take a cue from the Covid-19 crisis. To resolve the threat of the continuance of the virus, vulnerable communities were targeted to ensure that these people already on the fringes did not suffer any more than necessary. Government and health officials made strident efforts in the name of equity to resolve the occurrence of the virus in these edges of society. So, too, the proposal in this chapter has been to esteem the value of equity by addressing the digital needs of those on the peripheries of the global village with greater vigilance.

REFERENCES

Adekanmbi, G., & Boitshwarelo, B. (2012). Collaboration in Distance Education in Sub-Saharan Africa: Trends, Trials and. *Cases on Cultural Implications and Considerations in Online Learning*, 375.

Aisch, G. (2011). *Global digital divide*. Retrieved May 10, 2021, from http://old.driven-by-data.net/about/global-digital-divide/#/0

Alampay, E. (2006). Beyond access to ICTs: Measuring capabilities in the information society. *International Journal of Education and Development Using ICT, 2*(3).

Barzilai-Nahon, K. (2006). Gaps and bits: Conceptualizing measurements for digital divide/s. *The Information Society*, *22*(5), 269–278. doi:10.1080/01972240600903953

BBC News. (2013). *Kenya IT Hubs Launched for Primary Schools*. Retrieved May 10, 2021, from https://www.bbc.co.uk/news/world-africa-24147333

Bergquist, M., Ljungberg, J., & Rolandsson, B. (2011). A Historical Account of the Value of Free and Open Source Software: From Software Commune to Commercial Commons. In Open Source Systems: Grounding Research (pp. 196-207). Springer Berlin Heidelberg.

Carpentier, C. L., & Braun, H. (2020). Agenda 2030 for sustainable development: A powerful global framework. *Journal of the International Council for Small Business*, *1*(1), 14–23. doi:10.1080/26437015.2020.1714356

CBC News. (2011). *Internet 'kill switch' easy target in Egypt*. CBC News. https://www.cbc.ca/news/technology/internet-kill-switch-easy-target-in-egypt-1.1110730

Colle, R. D., & Roman, R. (2001). Editorial: The Telecenter Environment in 2002. *The Journal of Development Communication*, *12*(2), 15.

Compaine, B. M. (2001). *The digital divide: Facing a crisis or creating a myth?* Mit Press. doi:10.7551/mitpress/2419.001.0001

DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality. *Social Inequality: From Unequal Access to Differentiated Use*, 355-400.

Estes, A. (2011). The U.N. declares Internet access a human right. *National Journal.* Retrieved May 10, 2021, from https://www.nationaljournal.com/dailyfray/the-u-n-declares-internet-access-a-human-right-20110606

Evans, C. (2012). Five reasons we have yet to close the digital divide. *Microsoft in Education Blog*. Retrieved May 10, 2021, from, http://blogs.technet.com/b/microsoft_in_education/archive/2012/02/02/five-reasons-we-have-yet-to-close-the-digital-divide.aspx

Fink, C., & Kenny, C. J. (2003). W (h) ither the digital divide? *Info, 5*(6), 15-24.

Harmon, A. (1996). Daily Life's Digital Divide. *Los Angeles Times, 3*, A1. Retrieved May 10, 2021, from http://articles.latimes.com/1996-07-03/news/mn-20785_1_digital-technology

Hawkins, B. L., & Oblinge, D. G. (2006). The Myth about the Digital Divide. *EDUCAUSE Review*, *41*(4), 12–13.

Hilbert, M. (2010, August). *The manifold definitions of the digital divide and their diverse implications for policy responsibility*. TPRC.

Hilbert, M. (2011, December). Digital gender divide or technologically empowered women in developing countries? A typical case of lies, damned lies, and statistics. *Women's Studies International Forum*, *34*(6), 479–489. doi:10.1016/j.wsif.2011.07.001

Internet Live Stats. (2021). Retrieved May 10, 2021, from https://www.internetlivestats.com/internet-users/

Internet World Stats. (2021). *Usage and Population Statistics*. Retrieved May 10, 2021, from https://www.internetworldstats.com/stats.htm

Kularski, C., & Moller, S. (2012). The digital divide as a continuation of traditional systems of inequality. *Sociology*, *5151*, 1–23.

Lai, J., & Widmar, N. O. (2021). Revisiting the Digital Divide in the COVID-19 Era. *Applied Economic Perspectives and Policy*, *43*(1), 458–464. doi:10.1002/aepp.13104 PMID:33230409

Maguire, A. (2020). *The difference between the terms equality, equity, and liberation, illustrated*. Interaction Institute for Social Change.

Norris, P. (2000). The worldwide digital divide. In *Paper for the Annaul Meeting of the Political Studies Association of the UK*. London School of Economics and Political Science.

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge University Press. doi:10.1017/CBO9781139164887

Perrin, A. (2019). Retrieved May 10, 2021 from Digital Gap between Rural and Nonrural America Persists. https://www.pewresearch.org/fact-tank/2019/05/31/digital-gap-between-rural-and-nonrural-america-persists/

Poole, G. A. (1996, Jan. 29). A new gulf in American education, the digital divide. *New York Times*, p. 2.

Potter, A. B. (2006). Zones of silence: A framework beyond the digital divide. *First Monday*, *11*(5). Advance online publication. doi:10.5210/fm.v11i5.1327

Prieger, J., & Hu, W. (2006). An empirical analysis of indirect network effects in the home video game market. Retrieved May 10, 2021, from: http://ssrn.com/abstract=941223 doi:10.2139/ssrn.941223

Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, *6*(3), 341–362. doi:10.1177/1461444804042519

Simpson, L., Daws, L., & Pini, B. (2004). Public internet access revisited. *Telecommunications Policy*, *28*(3), 323–337. doi:10.1016/j.telpol.2003.10.001

Sparviero, S., & Ragnedda, M. (2021). Towards digital sustainability: the long journey to the sustainable development goals 2030. *Digital Policy, Regulation and Governance*.

Statista. (2021). Retrieved May 10, 2021 from www.statista.com/

United Nations Sustainable Development Goals. (n.d.). Retrieved May 10, 2021 from https://sdgs.un.org/goals/goal17

United States Department of Commerce's National Telecommunications and Information Administration. (n.d.). *Internet and Computer Use Studies and Data Files*. Retrieved May 10, 2021 from https://www.ntia.doc.gov/data

Van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, *19*(4), 315–326. doi:10.1080/01972240309487

Warschauer, M. (2002). Reconceptualizing the digital divide. *First Monday*, *7*(7). Advance online publication. doi:10.5210/fm.v7i7.967

Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT press. doi:10.7551/mitpress/6699.001.0001

Warschauer, M. (2012). The digital divide and social inclusion. *American Quarterly*, *6*(2), 131.

Wilson, J., & Wilson, H. (2009). Digital Divide: Impediment to ICT and Peace Building in Developing Countries. *American Communication Journal*, *11*(2), 1–9.

Yang, G. (2013). *The power of the Internet in China: Citizen activism online*. Columbia University Press.

Chapter 7
Ethical and Regulatory Challenges of Emerging Health Technologies

Samia Hassan Rizk
Faculty of Medicine, Cairo University, Egypt

ABSTRACT

The advances in biotechnology and computer and data sciences opened the way for innovative approaches to human healthcare. Meanwhile, they created many ethical and regulatory dilemmas such as pervasive global inequalities and security and risk to data privacy. The assessment of health technology is a systematic multidisciplinary process that aims to examine the benefits and risks associated with its use including medical, social, economic, and ethical impacts. It is used to inform policy and optimize decision-making. The advance of technology is creating significant challenges to healthcare regulators who strive to balance patient safety to fostering innovation. The FDA and EMA are modernizing their regulatory approaches to foster innovation in digital technology and improve safety and applicability to patients. On the other hand, data analytic technologies have been introduced into regulatory decision processes.

INTRODUCTION

The un-precedent power implied to humans with the aid of technology is transforming many of our life decisions, mostly for the better. Unless responsibly applied, the worse could also happen. The advances appearing in many biotechnology sciences such as genomics, neuroscience, synthetic biology, and nanoscience combined with the rapid developments in computer and data analytics allowed innovative approaches to human health care. On the other hand, the increasing complexity of ethical issues associated with modern health technology has created some novel challenges.

The need for specialized expertise and wider scope of analysis derived a new subfield of ethics termed technology ethics or "techno-ethics" which deals with the framing of principles and methods to guide technology implementation and use.

DOI: 10.4018/978-1-7998-8467-5.ch007

Common areas of techno-ethics include access right, intellectual property (IP) and privacy rights protection, environmental safety and reservation, human health and safety, technology predictability, transparency, and accountability (Spacey, 2016).

Health care is a major field of technology advancement involving a wide spectrum of issues such as medications, devices, innovative therapy, reproduction, personalized medicine, etc., and is associated with an increasing number of ethical dilemmas.

This chapter will review the main categories of health technologies and their potential ethical challenges, assessment methods, and related regulatory aspects.

BACKGROUND

The "Health Technology" term was defined in the 6Th World Health Assembly (2021) as: "the application of organized knowledge and skills within the sort of devices, medicines, vaccines, procedures, and systems developed to unravel ill health and improve quality of lives". (WHO, 2021)

The assessment of health technologies is a multidisciplinary process that systematically examines evidence about their various impacts including medical, social, economic, and ethical aspects. A comprehensive assessment requires knowledge and experience beyond the scope of ethical analysis including awareness of the spectrum and principles of the technology and the related impacts on all stakeholders. Transparency, freedom from bias, and robustness are essential qualities to ensure the validity of the assessment (www.eunethta.eu)

In the next sections, the spectrum of health technology, its ethical and regulatory challenges, assessment frameworks, and regulations will be discussed.

1. The Definition and Spectrum of Health Technology

The term 'health technology' encompasses a wide range of interventions, used for disease prevention, diagnosis, and treatment as well as rehabilitation and long-term care. In other words, it includes interventions intended to promote individual and population health.

It also applies to other broad applications such as restructuring of the health service through standardization or reallocation of resources.

Health technology can be categorized according to its type and intended use into:

1. Physical agents such as drugs and biologics e.g., vaccines and gene and cell products, devices and supplies e.g., cardiac pacemaker, implants, and magnetic resonance imaging (MRI) scanner, and procedures both medical and surgical
2. Public health programs e.g., newborn screening and immunization
3. Support Service e.g., clinical laboratory, blood bank, electronic health record system, telemedicine, and drug formulary
4. Organizational and managerial systems e.g., medication adherence program and alternative health care

Some applications combine more than one category e.g., drugs and devices (Lauritsen 2009) such as the positron emission tomography (PET) used with radiopharmaceuticals.

Innovative technologies can also be divided into evolutionary, the most common, and revolutionary, which is sporadically applied and include innovative applications in different subfields. (Iserson and Chiasson, 2002)

2. The Main Ethical and Regulatory Challenges of Emerging Health Technologies

Medical ethics are known since early civilization. However, advances over centuries brought up new concerns ranging from conception to end of life and beyond. Some of these may have legal aspects e.g., the informed consent, end-of-life decisions, and IP rights (Scaife, 2014). Several ethical dilemmas are evolving with the increasing use of health technology such as pervasive global inequalities, security, and privacy risks of data storage, sharing of databases, bio-repositories, bio-banks, etc.

Certain technologies have transformed the healthcare system particularly data-intensive medicine, omics, personalized medicine, and ‘remote’ forms of medicine such as e-health, m-health, telemedicine, and online health resources (The European Group on Ethics in Science and New Technologies, 2021). Their special impacts may be related to the perceptions of the ‘self, personhood, and body, the tensions between empowerment, engagement, and exploitation as well as the patient-physician relationship. Advanced technology will also affect societal understandings, principles, and structures governing health, and notions of solidarity and justice.

According to MacKay and Danis (2016); good governance, risk evaluation, and consideration of ethical values are important guiding elements in the selection and application of technology. However, a transformation of the regulatory system does not undergo a parallel transformation resulting in a widening oversight gap (Nash, 2021)

As reported recently by Credit Suisse Securities (2021), digital health adoption showed a “huge spike” during 2020 due to the increased need for virtual care, resulting in what is called a “great digital divide” particularly affecting; elderly population with lower digital ability in using patient portals (e.g., My Chart), lower-income populations and countries, and those with lower education levels. (Nash, 2021) The current COVID 19 pandemic has flared this digital divide by increasing digital technology applications such as the digital tracing for infected and exposed people using CCTV cameras, drones, credit card information, and location data from cell phones, with a privacy infringement.

In this situation, the restriction of some fundamental rights and freedoms may be ethically justifiable. However, limitations apply. Circumstances for restricting rights in these situations must be time-bound, meet standards of necessity, proportionality, and scientific validity. The technology used should conform to the privacy protection law and regulations, according to the principles of “The European Convention on Human Rights”, the “United Nations International Covenant on Civil and Political Rights”, and the “United Nations Siracusa Principles”. Thus, the fine line between saving lives and harming the fundamental rights and freedom of individuals should always be attended to even in critical situations.

For a good understanding of the ethical challenges of HT, a deep knowledge of the basic ethical principles and conduct criteria of the healthcare profession is called for.

Ethical Principles and Professional Conduct Criteria of Healthcare

The main conduct principles in medicine target the patient’s benefit in the first place and reflect the 4 foundational ethical principles of (1) autonomy, (2) beneficence, (3 BM) non-maleficence, and (4) justice.

As told in the American Medical Association (AMA) permeable (1995), a physician has a responsibility towards, patients, society, other health professionals, and to self. (AMA,1995).

Nine professional Principles are included (Beauchamp and Childress, 2013):

1. Dedication to providing competent medical care, with compassion and respect for human dignity and rights.
2. Upholding the standards of professionalism, honesty in all professional interactions, and striving to report physician deficient in character or competence, or engaging in fraud or deception, to proper entities.
3. Respecting the law and recognizing a responsibility to seek changes in those requirements which are contrary to the best interests of the patient.
4. Respecting the rights of patients, colleagues, and other health professionals, and safeguarding patient confidence and privacy within the constraints of the law.
5. Continuing to study, apply, and advance scientific knowledge, keeping a commitment to medical education, making relevant information available to patients, colleagues, and the public, obtaining consultation, and using the talents of other health professionals when indicated.
6. In the provision of proper patient care, except in emergencies, be free to choose whom to serve, with whom to associate, and the environment in which to supply medical care.
7. Recognizing a responsibility to take part in activities contributing to the improvement of the community and the benefit of public health.
8. While caring for a patient, regard responsibility to the patient as paramount.
9. Supporting access to medical care for all people.

In certain situations, however, novel challenges may conflict with any of the above principles, whether in day-to-day or policymaking decisions. The commonest of these challenges include (Advent Health, 2021):

- The increasing use of big data and stored biological samples for predictive analytics, and algorithms. These can lead to patients' exploitation by not obtaining their consent, privacy disclosure, and commercialization of data and samples.
- Collection of population health data through public health programs e.g., newborn screening without parental permission.
- The digitalization of health records with accessibility to privacy violation. Although techniques for data security are implemented, the protection is not absolute and re-identification can still occur.
- The use of customized therapeutics and diagnostics based on personal genomic data may enhance inequality.
- Increased technology costs will compromise the balance between patients' needs and fiscal responsibility and could unintentionally impose on patients' best interests. Such conflicting responsibilities could lead to "moral injury", particularly in a profit-driven health care environment. (VILARDELL, 1990)
- Applications of Artificial Intelligence (AI) for decision-making in many healthcare fields constitute a potential source of abuse when applied for other purposes such as income or criminal data. These actions are prohibited by the "Genetic Information Nondiscrimination Act" of 2008

- Robotics used as an assist tool in surgery and other uses in the future such as care for the elderly to compensate for the shortage of healthcare workers which level might be dehumanizing.

At the institutional level, research ethics committees are facing challenges with the review of big data research including scientific and legal issues in judging data identifiability thresholds and security risk as well as the revision of contractual obligations as the DUAs and MOUs. Another important challenge is the consent process and commitments of the researcher. With retrospective data research, waiving of consent for the use of anonymized data and or biologic samples, though practically justified may allow an unauthorized use in the new technological era. With prospective studies the traditional consent models often needs detailed description of the extent of confidentiality of data which might not be understood by the participant. (Draper & Owen, 2016)

According to their cost-benefit effects medical technologies have been classified by the Royal College of Surgeons into three categories: those applied for disease prevention and health promotion with little expenditure; those used for treatment with moderate cost; and those used to support health and quality of life but with high expenditure. This last category is considered the most challenging (Black, 1979).

Another impact of technology in the 21st century is a change in basic assumptions in the healthcare system encouraging citizens' involvement raising new ethical consequences including a shift in the belief, organization, and delivery of healthcare, destabilization of traditional structures of power, and knowledge unpinning medical practice. Risks could arise when technologies cause patients to lose critical contact with medical professionals. The EGE draws some recommendations in response to these changes including the balancing autonomy and responsibility as well as considerations for justice and solidarity in which decisions about expensive, high-tech treatments should be carefully balanced with the wider social needs of healthcare.

The increased capabilities of what can technology do in human life raise ethical debates about the concept of life and death. The advancement in stem cell research, cloning, and gene editing are just a few examples of debates about the essential definition of "human" and the ethical responsibilities of biological scientists. (Scaife, 2014). The Slowing down of aging and increased life expectancy would also create concerns such as who makes the rules about how long people should live? And how to afford these longer lives and choose who gets access to health care? Similarly, they could raise questions about our ability to manipulate and dehumanize individual life. The National Academy of Sciences (year) developed guidelines for the control of genome editing which prohibit any use of the technology to create "designer babies" or for "human enhancement."

Some innovative approaches in information technology that are integrated with the human body (ICT implants) such as the brain-computer interfaces in neuroscience could potentially alter a patient's personality traits which is still an open debate. (Marckmann and Goodman, 2006). With life-sustaining technologies, doctors and families must make complex decisions about when to "pull the plug". These decisions involve discussions about the quality of life as well as the financial resources. Another evolving difficulty in this issue is the decision about euthanasia whether should it be legalized. (Sirico, 1997)

Beyond the ethical issues, some legal challenges may also be relevant to emerge health technologies. An important example is the patenting rules. While important for ensuring funding and providing incentives for producers particularly in areas of limited applications, patenting is associated with some ethical and legal issues including; increased costs with resulting health inequality and injustice, inhibiting further research and innovation by denying access to genes or other biological material and creating an "anti-commons" effect, where scientists avoid avenues of research that are long and expensive. Besides, patenting of human biologic material is condemned by some parties as a disrespectful act. (Cook-Deegan, 2008). Another legal issue is the law enforcement access to personal or family genetic information which may affect individuals' willingness to undergo health care testing or take part in research. (Hazel and Clayton, 2021) Recent initiatives were developed to overcome these acts such as condemning the indefinite retention of bio-specimens and data by the "European Court of Human Rights". In the U.S., the third-party genetic information sources allow individuals to choose whether they want their submitted data used for law enforcement. A related issue is the use of advanced technology like DNA phenotyping to generate virtual sketches of suspects in police stations which is a violation of privacy. (Hazel and Clayton, 2021)

3. The Assessment of Health Technology (HTA)

The HTA involves systematic evaluation of the properties and application consequences of the technology, both direct (intended) and indirect (unintended). It is used in many ways to inform healthcare decisions such as permitting commercial use of a product by regulatory agents, payment plans, drug formularies, insurers, clinicians, and patients; development of practice guidelines; technology acquisition and management by hospitals or other health care organizations; specify technology standards, public health programs, policies for technology innovation, research, and funding, product development and marketing decisions by companies and other investors, detecting gaps and unmet health needs by research agencies.

The term "technology assessment" was introduced in 1965 by the "Committee on Science and Astronautics of the US House of Representatives". It was defined 30 years ago by the WHO to strengthen evidence-based choice and rational use of health technologies and was later used as a decision tool for the implementation of Universal Health Coverage (UHC) (Carrin et.al. 2007). Recently, a new definition was introduced by the "International Network of Agencies for Health Technology Assessment (INAHTA)" and the "Health Technology Assessment International (HTAi)" stating that: "HTA is a multidisciplinary process that uses explicit methods to determine the value of health technology at different points in its lifecycle. The purpose is to tell decision-making to market a fair, efficient, and high-quality health system". (O'Rourke et.al. 2020)

The Concept and Methods of HTA: The framing, methods, and decisions used in HTA reflect the values of various stakeholders as well as the consequences of technology implementation (Legault et.al. 2019).

The methods for HTA were first described by Banta in 1976 (Banta et.al. 2009) and were further developed by the "National Academy of Sciences", the "National Academy of Engineering (NAE)", and the "Legislative Reference Service of the Library of Congress".

The concept of TA embraces different methods of policy analysis and has three dimensions: Cognitive (knowledge); Normative (dialogue); and Pragmatic (processes). Three aspects of assessment are involved: the issue or the technology; the social aspect; and the policy aspect (European Parliamentary Technology Assessment (EPTA), 2013).

The HTA relies on "systems theory" which assumes that complex ideas can be viewed as systems with common designs and properties which can be studied using systems' method. These systems are dynamic with changing relations between technology, nature, and society.

Two main aspects of technology are examined; the Techno-ethical design (TED) where designers acknowledge properties of the technology concerning its intended use. The other is the Techno-ethical assessment (TEA) which examines the knowledge, goals, inputs, and outputs of the system.

The important aspects of HTA may vary with the aim. However, the essential ones defined by the WHO global survey on HTA (2015) include safety, clinical effectiveness, equity, ethical issues, acceptability to health care providers, and acceptability to patients.

The HTA and its Optimal Use (OU) starts with topic suggestions from various sources; individuals, or entities. The topics are prioritized according to predetermined scoring criteria and then a comprehensive evaluation of the clinical, economic, and implementation aspects is given including ethical, legal, and social issues. Final reporting with recommendations for OU and HTA is then developed. (CADTH, 2015)

The conceptual foundation for technology analysis was developed into interdisciplinary frameworks with a systems-based approach that involves five key steps including:

1. The intended use and side effects, reflecting its overall value (interest).
2. The intended use is compared to the alternatives or the best standard of care in terms of technical and non-technical (moral and social) aspects.
3. Excluding non-conforming overall value in terms of efficiency and fairness.
4. Considering the stakeholders' perspectives.
5. Evaluating various impacts of the technology e.g., biological, physical, psychological, social, and environmental.

Given its wide scope, a range of expertise is often needed for HTA e.g., physicians and other healthcare workers, managers, biomedical scientists, epidemiologists, community representatives, lawyers, computer scientists, in addition to ethicists. The choice of the ability will depend on the scope of the technology.

Diverse methods were developed by different sectors including; for-profit and not-for-profit private sectors and government agencies to assess technologies at their various stages of the lifecycle, including. (NICHSR, 2021):

- Future: during a conceptual stage, expected, or within the earliest stages of development
- Experimental: during bench or laboratory testing
- Investigational: during first clinical (i.e., in humans) evaluation for a particular condition or sign
- Established: considered by clinicians as a standard approach and applied into widespread use
- Obsolete/outdated/abandoned.

However, technologies may not mature linearly through these stages e.g., an established product for certain indications investigated for another, reintroduction of a previously obsolete technology for a new indication such as the drug thalidomide (Breitkreutz et.al. 2008; Zhou et.al. 2013).

Three main approaches for HTA are described with overlap and combination, including:

1. Technology-oriented approach: intended to figure out the characteristics or impacts of technologies.
2. Problem-oriented approach: focuses on solutions or strategies for managing a particular disease or condition
3. Project-oriented approach: involves the local placement or use of technology in a particular institution, program, or project.

Two of the main types of HTA methods include primary data collection and secondary or integrative methods. The first involves the collection of original data from many studies. The second which is more commonly used involves integrating data from pre-existing sources. All data and information used are checked for quality and risk of bias. (HIQA, 2016)

A recent trend in HTA is the cooperation of different counties and organizations in decision-making through common platforms such as the EUnet HTA platform across Europe. This approach has the advantages of identification of common unmet needs, support of assessment and decision-making processes, and sharing of resources. (https://www.medtecheurope.org/) A "core model" for HTA was developed by the EUnet as a generic framework to enable international collaboration (EUnet, 2013).

The model defines 7 phases and 9 domains for a comprehensive HTA (EUnet HTA 2008). These include:

Core HTA model phases:

1. Definition of the technology to be assessed
2. Definition of project type
3. Relevance of assessment elements
4. Translation of relevant issues into research questions
5. Compiling of a core HTA protocol
6. Research
7. Entering the results

Core HTA model domains:

1. Defining health problem and the current use of technology
2. Description of the technical characteristics of technology
3. Safety
4. Clinical effectiveness
5. Costs and economic evaluation
6. Ethical analysis
7. Organizational aspects
8. Social aspects
9. Legal aspects

Certain technologies such as AI applications may require specific assessment methods. Guidelines for developing, using, and regulating AI were issued by the American Medical Association (AMA) in 2018. With the increasing applications of AI, special guidance for their evaluation was developed by the "EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES" and included 8 ethical principles related to safety, avoidance of bias, transparency, privacy and data protection, decision making, liability, human values, and governance. (Osman, 2019)

The adoption of the "Health Technology Assessment and incorporation into Health Systems" resolution has been introduced by the Pan American Health Organization (PAHO) (CSP28.R9) (PAHO, 2012). Another initiative is the Regional Platform on Access and Innovation for Health Technologies (PRAIS) which helps linkages between stakeholders from a public health perspective. Other international and regional HTA networks were developed to share information and best practices, and build capacity. (WHO, 2021)

An emerging aspect of HTA is patients' engagement to improve decision making through improving quality and enhance data collection. An example is the "The international alliance of patients' organizations" (https://www.iapo.org.uk/ (2017). Guidance on patient involvement in HTA with actionable recommendations on working practices was developed by The EUPATI HTA (Hunter et.al. 2018).

The involvement of pharmacists as important stakeholders in HTA has been highlighted by the Pharmaceutical Group of the European Union (PGEU) in their position paper on big data and AI in healthcare. The group acknowledged the impact of this technology in promoting safety and rationale use of medicinal products by community pharmacies. (PGEU, 2016)

The development of Technology Ethics Committees by healthcare organizations as an added review body for emerging issues related to technology applications was suggested by Cossitt (2020)

1. The Current Regulatory and Legislative Status of Health Technology

Traditionally, health technologies are regulated within medical devices which are reviewed through an appropriate premarket pathway, like premarket clearance (510(k)) (/medical-devices/premarket-submissions/premarket notification- 510k) (FDA, 2020), De novo classification(/medical-devices/premarket-submissions/de-novoclassificationrequest) (FDA, 2019), or premarket approval. (/medical-devices/premarket submissions/ premarket-approval-pma) (FDA, 2019). Obscure reporting with limited data on harm and malfunctions for medical devices were allowed by the FDA (Kaplan, 2021

This system, however, is not compatible with recent digital applications such as adaptive artificial intelligence and machine learning technologies. Due to the rapid evolution and increased applications of these devices, ethical concerns have been raised e.g., hacking of devices like pacemakers, violation of privacy, increasing disparities in care, and rising costs compared to effective alternatives (UIC, 2020)

The first FDA regulatory guideline (1976) of health technology defines it as "an instrument, apparatus, implement, machine, contrivance, implant, in vitro substance, or another similar or related article, including any component, part, or accessory…" With the growth of the health technology market, a simplified approach was developed for applications such as wearable but with no binding documents. The "Med Tech Act" introduced in 2014, exempted electronic health records and other consumer health software from FDA regulation. The "SOFTWARE Act" (Sensible Oversight for Technology which Advances Regulatory Efficiency) then indicated non-regulation of any device that did not pose a serious health risk to the consumer if the device malfunctioned or those used for fitness or basic health metrics tracking.

To streamline approval of regulated devices the FDA introduced the "Health Technology Regulations" to replace Group III products within the "Hazardous Substances Act" of 1973 Classification of medical devices. Licensing medical devices classifies them into 4 risk groups: I Low-Risk, IIa Low-moderate Risk, IIb Moderate-high Risk, and III High Risk. (EGGERS et al. 2018)

The Regulatory Challenges of Health Technology

Health regulators strive to balance patient safety and the benefits of technological innovations. Digital technologies are particularly challenging especially with the new business models, which require the modification, enforcement, and communication of regulations while avoiding overregulation.

Two categories of technology challenges are described:

1. Technological challenges including data, and digital privacy, and security issues as well as AI-based issues. Globally digital security is guided by different laws. In the EU, the general data protection regulations (GDPR) force strict control over data transmission. While in the US they focus on special sectors such as healthcare. An example of the latter is the FDA rule for the software "SaMD" as a medical device, an assignment that is not compatible with its development and validation. Another challenge to digital innovation is cyber-security, where digital health devices are continually collecting and analyzing data. The AI-based technology presents 2 main specific problems including algorithmic bias, which is inherent with some AI applications. The other is the "black box" problem where algorithms are often held by their creators and are so complex for the user, leading to the inability to know what is inside. This latter problem may be related to the non-disclosure agreements. Recently, the EU GDPR (2018) mandated a clear explanation of data-driven decision devices by companies.
2. Business challenges include the pacing problem describing the gap between the rapid technology advancement and the existing regulatory mechanisms which are getting wider and further complicated by the regulatory patchwork. Another challenge is the disruptive business models with interconnected nature, making it difficult to assign liability, and requiring more than one regulatory application. Furthermore, the limitation of research on regulatory aspects such as empirical investigation of the impacts of alternative regulatory approaches on important consequences as health outcomes is delaying regulatory innovations (Kimbrell, 2018)

To face these challenges new regulatory approaches have been proposed including:

1. Adaptive approach: a responsive iterative method based on trial and error with co-design of regulations and standards and rapid feedback loops. It contrasts with the traditional "regulate and forget" one.
2. Regulatory sandboxes: involve design and test of a product prototype and involvement of partners with a spectrum of expertise from diverse sources to accelerate regulatory design
3. Outcome-based approach: focuses on outcome performance measures
4. Risk-weighted approach: is data-driven and involves the collection and advanced data analysis thus allowing accurate, safe, effective, and personalized technology product. This contrasts with the old concept of "one size fits all" and has been recently adopted by the FDA for certain digital health products.
5. The "PRE-CERT PILOT PROGRAM" is a part of the "DIGITAL HEALTH INNOVATION ACTION PLAN", intended to accelerate approval of low-risk products through a simple premarket review.
6. Collaborative approach: aligns regulatory views by engaging all stakeholders across the echo-system into policy guidance and standards.

For any proposed regulatory transition, it is important to consider some foundational aspects such as the current regulatory status i.e., pre-existing regulation which might be blocking innovation, the right time for a change, and the suitable approach, with consideration of what is called" global innovation arbitrage" (Thierer, 2019). It is also important to simplify the regulatory processes and minimize its phases as indicated by the European Medicines Agency (2020)

Digital technologies are increasingly used in clinical trials e.g., for patient monitoring, electronic data capture, and electronic consent signatures, with questionable reliability for validation of benefit-risk data analysis, compared to the established GCP requirements. Currently, the EMA accepts the specific use of these technologies in the development, evaluation, and use of medicines including sensors, mobile health tools, and telehealth as well as health data analytics and digital record systems in clinical trials (EMA, 2020), where any potential impact on the benefit-risk assessment of "Marketing Authorization Application" (MAA) is discussed at the time of evaluation. Examples of digital methodologies qualification programs include Digital endpoint; Digital Biomarker (BM), Electronic Clinical Outcome Assessment (eCOA), and Digital measures. Any added requirements throughout product development should also be considered e.g., device legislations, ICH E6 (R2) GCP, and ICH guidance.

Thus, the focus of the EMA qualification is to ensure the reliability of the technology-measured clinical parameter for the intended use during drug development, meaningful interpretation of the concept of interest, and robustness and reliability of the method used to include the context (s) of use (CoU) which are often provided by pivotal risk-benefit data assessment. Furthermore, the nature and impact of any change in the product lifecycle should be considered. (Corvus et al., 2020 and Gottlieb, 2018)

On the other hand, the simultaneous implementation of a follow-up mechanism is important to keep relevant regulations. This can be achieved through a planned revision scheme or including tools for automatic review such as sunsetting applied by the "European Union's regulatory fitness and performance (REFIT) program" (Wiener, 2004).

Governance is another aspect of regulatory supervision. Over the past three decades, it has been achieved through assessment to forecast the social and economic impacts of innovative technologies, and those of regulations (regulatory impact assessments (RIAs)). Initiatives for improving governance, guidelines, data quality framework, and infrastructure are supported by the European Health Data Space (EHDS) through the promotion of health data exchange, support of health research and innovation as well as regulators access to health data for policymaking and investigating regulatory gaps. (Hazel and Clayton, 2021)

To streamline the path for digital health products, the FDA focused its oversight review on the safety and effectiveness of the higher-risk medical devices. The new Pre-Cert program allows the use of adaptive machine learning as well as the introduction of minor changes without resubmission. The FDA's (CDRH) total product lifecycle-based regulatory framework allows modifications to be made from real-world learning and adaptation while ensuring the safety and effectiveness of the software. This approach enables both the FDA and manufacturers to check products from premarket development to post-market performance (FDA, 2018). With this seamless passage 51, digital health products have been authorized in 2017.

While updating regulatory approaches, the FDA is also fostering the internal application of digital health tools and analytics e.g., using digital biomarkers and data from EHRs to enable pragmatic clinical trials and the "Premarket Digital Safety Program" which allows unification of data standards of electronic reporting systems under the expedited safety-reporting regulations. The later program was first applied in the FDA's "Oncology Center of Excellence", and the agency's drugs and biologics centers. Two further improvements are the "Information Exchange and Data Transformation" or INFORMED which supports the integration of data analytics into regulatory decision-making (Khozin, 2017), and the "Digital Health Innovation Action Plan" which improves the clarity and efficiency of digital health products regulation (Gottlieb, 2018)

A variety of regulatory science tools have been developed by the FDA's Center for Devices and Radiological Health's (CDRH) Office of Science and Engineering Labs (OSEL) to expand the scope of innovative science-based approaches and improve the development and assessment of emerging medical technologies. A catalog of more than 35 Regulatory Science Tools in different medical fields was developed. These tools are classified into 3 main types; Phantoms, Methods, and Computational Models and Simulations with several examples including; Medical imaging and diagnostics, CV and Neurology simulation models, digital pathology, Ultrasound, electromagnetic and electric safety, artificial intelligence, and machine learning models, orthopedic and ophthalmology therapeutics, material performance, etc. These tools, however, do not replace FDA-recognized standards or qualified medical device development tools (MDDTs) (FDA, 2021).

Until now five main actions promoted the use of AI/ML technology for oversight of SaMD and the field in general including (Mahler et.al. 2021)

1. Tailored Regulatory Framework for AI/ML-based SaMD: a new Draft Guidance on the Predetermined Change Control Plan of the proposed framework is developed.
2. Good Machine Learning Practice (GMLP): the importance of GMLP is supported by stakeholders who also call for harmonization of its development through consensus standards and other community initiatives. It helps products' oversight as well as the robustness of cybersecurity, in collaboration with the agency's "Medical Device Cybersecurity Program".
3. Patient-Centered Approach Incorporating Transparency to Users: this proactive approach is based on the unique properties of AI/ML devices such as usability, equity, trust, and accountability, addressed through transparency to users.
4. Regulatory Science Methods Related to Algorithm Bias & Robustness: improved methods for evaluation and addressing algorithmic bias and promoting robustness are introduced.
5. Real-World Performance (RWP): a collection of performance data on the real-world use of the SaMD will improve risk mitigation and enhance their improvements.

SOLUTIONS AND RECOMMENDATIONS

With the worldwide expansion of universal healthcare coverage HTA becomes an essential demand.

However, countries vary widely in their implementation of HTA, depending on many actors such as level of investment in health, political support, and infrastructure of the healthcare system. Rich countries for example have established HTA system with well-developed infrastructure and framework procedures. While low-economy countries are still striving to build new systems, especially with the rising healthcare costs.

The collaboration at regional and international levels will help to economize resources, transfer models, exchange expertise, and harmonize frameworks, while keeping the unique requirements of each country.

A well-developed system should integrate infrastructural elements, frameworks for assessment and governance, and procedural steps. Among the infrastructural elements, building expertise and assessment-workforce through educational or training programs is essential for a sustained capacity. Introducing innovative technology e.g., AI and machine learning will support HTA infrastructure and ensures its continuous development. The input from various stakeholders including patients will support and improve HTA based decisions.

FUTURE RESEARCH DIRECTIONS

The emerging field of HTA has many opportunities for innovative research to set up an integrated evidence-based system. exploring infrastructure needs specific for the country, and knowledge gaps is open to empirical research examining and comparing different governance and feedback methods are particularly important to ensure a continuously updating regulatory system. the research-driven decisions will help to set up a more fitting and responsive system to the country needs.

CONCLUSION

Health technology is rapidly evolving and transforming the healthcare system. Data-driven systems are particularly changing the traditional healthcare delivery. The gap in knowledge and expertise between these innovative interventions and the existing healthcare contexts for both givers and users, creates several ethical, regulatory, and economic challenges which could destabilize trust in the safety, fairness, and effectiveness of the healthcare system.

In addition, healthcare financing as an integral part of any country economy should be strongly based on accurate cost-effectiveness analysis particularly in countries with limited resources.

Hence the necessity of HTA to set up a reliable healthcare system which is responsive to the medical, societal, governmental, and global needs

REFERENCES

Advent Health Medicare Advantage. (2021). *Advent Health Medicare Advantage*. healthgolds.com

AMA Principles of Medical Ethics. (1995). *American Medical Association*. https://www.ama-assn.org/about/publications-newsletters/ama-principles-medical-ethics

Banta, D., Kristensen, F. B., & Jonsson, E. (2009). A history of health technology assessment at the European level. *International Journal of Technology Assessment in Health Care*, *25*(S1), 68–73. Advance online publication. doi:10.1017/S0266462309090448 PMID:19534837

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of Biomedical Ethics* (8th ed.). Oxford University Press.

Black, D. (1979). The paradox of medical care. *JR Co1 Physicians Land, 13*, 57-65.

Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H., Bähler, J., & Wood, V. (2008). The BioGRID Interaction Database: 2008 update. *Nucleic Acids Research, 36*(1), D637–D640, doi:10.1093/nar/gkm1001

Carrin, G., Evans, D., & Xu, K. (2007, September). Designing health financing policy towards universal coverage. *Bulletin of the World Health Organization*, *85*(9), 649–732. doi:10.2471/BLT.07.046664 PMID:18026615

Coravos, A., Doerr, M., Goldsack, J., & Wood, W. A. (2020). Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. *NPJ Digital Medicine, 3*(1). https://www.researchgate.net/deref/http%3A%2F%2Fwww.nature.com%2Fnpjdigitalmed

Cossitt, A. (2020). *Why Health Care Organizations Need Technology Ethics Committees. February 5, in Ethics*. Hastings Bioethics Forum, Health, And Health Care. https://www.thehastingscenter.org/

Draper, B. L., & Owen, S. (2016). *Big Data Research: Practical Solutions to Emerging Challenges for IRBs.* Webinar of the Public Responsibility in Medicine and Research (PRIM&R) organization. www.PRIM&R.org

Eggers, W.D., Torley, M., & Kishnani, P. (2018). *Principles of regulating emerging technologies*. A report from the Deloitte center for government insights. the future of regulation. https://www2.deloitte.com/us/en/insights.html

EUnetHTA. (2013). *The HTA Core Model is a methodological framework for shared production and sharing of HTA information*. The HTA Core Model is a registered trademark. Joint Action 2, Work Package 8. HTA Core Model ® version 2.0. http://www.corehta.info/BrowseModel.aspx

EUnetHTA. (n.d.). *Creating, Facilitating, and Promoting Sustainable Health Technology Assessment (HTA) Cooperation in Europe News_FF9900-500*. Retrieved March 29, 2021, from https://www.eunethta.eu

EUnetHTA Project. (2008). *Encyclopedia of Public Health* (W. Kirch, Ed.). Springer. doi:10.1007/978-1-4020-5614-7_1066

EUR-Lex. (2018). *The general data protection regulation applies in all Member States from 25 May 2018*. https://eur-lex.europa.eu/

European Group on Ethics in Science and New Technologies (EGE). (2021). https://ec.europa.eu/info/research-and-innovation/strategy/support-policy-making/scientific-support-eu-policies/

European Medicines Agency (EMA). (2020). *Qualification of digital technology-based methodologies to support approval of medicinal products*. EMA/219860/2020. https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance/qualification-novelmethodologies-

FDA. (1976). Medical Device Amendments of 1976. In *History of Federal Regulation: 1902–Present. Major legislation with regard to drugs and medical devices*. www.fdareview.org/

FDA. (2018). *Software as a medical device. An agile model for food and drug administration (FDA)-regulated software in health care*. https://www2.deloitte.com/

FDA. (2019). *Premarket Approval (PMA). 05/16*. https://www.fda.gov/medical-devices/

FDA. (2019). *De Novo Classification Request 11/20*. https://www.fda.gov/medical-devices

FDA. (2020). *Premarket Notification 510(k) 03/13*. https://www.fda.gov/medical-devices

FDA. (2021). *Catalog of Regulatory Science Tools to Help Assess New Medical Devices. 03/23*. https://www.fda.gov/medical-devices

Gottlieb, S. (2018). *Digital Health Innovation Action Plan*. https://www.fda.gov/media/106331/

Gottlieb, S. (2018). *Transforming FDA's Approach to Digital Health*. https://www.fda.gov/news-events/speeches-fda-officials/transforming-fdas-approach-digital-health-04262018

NASA authorization 1965. hearings before the Committee on Science and Astronautics, U.S. House of Representatives, Eighty-eighth Congress, second session, on H. R. 9641, superseded by H. R. 10456. Washington:GovtU. S. https://catalog.hathitrust.org/Record/100718937/Cite

Hazel, J. W., & Clayton, E. W. (2021). *Law Enforcement and Genetic Data*. Bioethics Briefings. The Hasting Center. https://www.thehastingscenter.org/briefingbook

Iserson, K.V., & Chiasson, P.M. (2002). The ethics of applying new medical technologies. *Semin Laparosc Surg., 9*(4), 222-9. doi:10.1053/slas.2002.36465

Kaplan, B. (2021). *Regulation of Software as a Medical Device: Opportunity for Bioethics*. Hastings Bioethics Forum. https://www.thehastingscenter.org/

Khozin, S. (2017). Information Exchange and Data Transformation (INFORMED) An integrated approach to big data analytics. *Nat Rev Drug Discov*. https://www.ehidc.org/

Kimbrell, G. (2018). *Is Regulation Killing Innovation in Health Care?* Forbes Technology Council. https://www.forbes.com/sites/forbestechcouncil

Lauritsen, K. J., & Nguyen, T. (2009). Combination Products Regulation at the FDA. *Clinical Pharmacology and Therapeutics*, *85*(5), 468–470. doi:10.1038/clpt.2009.28 PMID:19381151

Legault, G. A., Béland, J. P., Parent, M., Bédard, S. K., Bellemare, C. A., Bernier, L., Dagenais, P., Daniel, C. E., Gagnon, H., & Patenaude, J. (2019). Ethical Evaluation in Health Technology Assessment: A Challenge for Applied Philosophy. *Open Journal of Philosophy*, *9*, 331–351. doi:10.4236/ojpp.2019.93022

MacKay, D., & Danis, M. (2016). Federalism and Responsibility for Health Care. *Public Affairs Quarterly*, *30*(1), 1–29.

Mahler, M., Auza, C., Albesa, R., Melus, C., & Wu, J. A. (2021). Regulatory aspects of artificial intelligence and machine learning-enabled software as medical devices (SaMD). doi:10.1016/B978-0-12-820239-5.00010-

Marckmann, G., & Goodman, K. (2006). Introduction: Ethics of Information Technology in Health Care. *International Journal of Information Ethics*, *5*, 2–5. doi:10.29173/irie188

Nash, D. (2021). *Healthcare Technology's Digital Dilemma. Should digital connectivity be considered a "vital sign"?* https://www.medpagetoday.com

National Information Center on Health Services Research & Health Care Technology (NICHSR). (2021). *National Library of Medicine*. https://www.nlm.nih.gov/

O'Rourke, B., Oortwijn, W., & Schuller, T.International Joint Task Group. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187–190. doi:10.1017/S0266462320000215 PMID:32398176

Osman, H. (2019). *SA Health to overhaul EPAS. Healthcare IT news*. https://www.healthcareit.com.au/

Pan American Health Organization World Health Organization. (2012). *Health Technology Assessment and Incorporation into Health Systems*. RESOLUTION CSP28.R9 Sixth meeting, 19 September. https://iris.paho.org/handle/10665.2/3684

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., Gordon, M., Hermanson, L., Knight, J. R., Lea, D. J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A. K., Smith, D., Vellinga, M., ... Williams, A. (2014). Skillful Long-Range Prediction of European and North American Winters. Met Office Hadley Center. *Geophysical Research Letters*, *41*(7), 2514–2519. doi:10.1002/2014GL059637

Sirico, R. A. (1997). *The Religious Left with Its Mask Off*. https://www.acton.org/

Spacey, J. (2016). *20 Types of Technology Ethics*. Simplicable. https://simplicable.com/

The American Medical Association (AMA). (2018). *AMA adopts a policy on augmented intelligence*. https://medicalxpress.com/news

The Canadian Agency for Drugs and Technologies in Health CADTH. (2015). *Health Technology Expert Review Panel: Process for Developing Recommendations.* Retrieved from https://www.cadth.ca/sites/default/files/pdf/HTERP_Process.pdf

The European Coordination Committee of the Radiological, Electromedical and Healthcare IT Industry (COCIR). (2019). *European Health Data Space: Towards a Better Patient Outcome*. https://www.cocir.org/fileadmin/Publications

The European Medicines Agency's (EMA). (2020). *Regulatory science strategy*. https://www.ema.europa.eu/

The International Alliance of Patients' Organizations (IAPO). (2017). https://www.iapo.org.uk

The Pharmaceutical Group of the European Union (PGEU). (2016). *Position P aper on Big Data & Artificial Intelligence in Healthcare.* Ref 19.02.20E 001. https://www.pgeu.eu/

Thierer, A. (2019). *Global Innovation Arbitrage: Export Controls Edition*. The Technology Liberation Front. https://techliberation.com/

University of Illinois. (2020). *Emerging Technologies and Their Impact on Health Informatics*. https://healthinformatics.uic.edu/

Vilardell, F. (1990). Ethical problems of medical technology. *Bulletin of the Pan American Health Organization*, *24*(4), 379–385. PMID:2073552

Wiener, J. B. (2004). The regulation of technology, and the technology of regulation. *Technology in Society*, *26*(2-3), 483–500. doi:10.1016/j.techsoc.2004.01.033

World Health Organization (WHO). (2015). *Global Survey on Health Technology Assessment by National Authorities.* www.who.int/health-technology-assessment

World Health Organization (WHO). (2021). https://www.who.int/

Zhou, Y., Chen, C., & Johansson, M.J.O. (2013). The pre-mRNA retention and splicing complex controls tRNA maturation by promoting TAN1 expression. *Nucleic Acids Research, 41*(11), 5669–5678. doi:10.1093/nar/gkt269

ADDITIONAL READING

Hasan, N., & Bao, Y. (2021). Understanding current states of machine learning approaches in medical informatics: A systematic literature review. *Health Technology (Hong Kong)*, *11*(3), 471–482. doi:10.100712553-021-00538-6

Kaiser, F. K., Wiens, M., & Schultmann, F. (2021). Use of digital healthcare solutions for care delivery during a pandemic-chances and (cyber) risks referring to the example of the COVID-19 pandemic. *Health Technology (Hong Kong)*, *11*(5), 1125–1137. Advance online publication. doi:10.100712553-021-00541-x PMID:33875933

Kalaiselvan, V., Sharma, A., & Gupta, S. K. (2021). "Feasibility test and application of AI in healthcare"—With special emphasis in clinical, pharmacovigilance, and regulatory practices. *Health Technology (Hong Kong)*, *11*(1), 1–15. doi:10.100712553-020-00495-6

Novaes, H. M. D., & Soárez, P. C. (2016). Health technology assessment (HTA) organizations: dimensions of the institutional and political framework. *Cad. Saúde Pública, Rio de Janeiro, 32* Sup 2: e00022315, 2016. doi:10.1590/0102-311X00022315

KEY TERMS AND DEFINITIONS

Beneficence: Is the act of benefitting others which is an obligation of the physician towards his patient and is one of the bioethics principles.

Cyber Security: Is the mechanism of protecting computer systems and networks against exploitation thus reducing risk of disclosure of confidential data.

Decision Algorithms: Is a hierarchical flow of answers to an upper-level question based on the input data for a given decision problem.

Gene Editing: Is the change of a gene constitution by inserting external DNA pieces with the use of technology. It has a therapeutic benefit in some genetic disorders.

Governance: Is the act of supervised control of a given working party by an authorized body.

Machine Learning: Is a computer algorithm program which automatically learns and adapts to new data without human interference.

Techno-Ethics: Is the subfield of ethics which analyzes the addresses the ethical dimensions in the production and application of technology.

Tele-Health: Is the delivery of healthcare service or communication on a distance via technology.

Chapter 8
Ethical Benefits and Drawbacks of Digitally Informed Consent

Wendy Charles
https://orcid.org/0000-0003-1627-6899
BurstIQ, USA

Ruth Magtanong
Case Western Reserve University, USA

ABSTRACT

As organizations steadily adopt remote and virtual capabilities, informed consent processes are increasingly managed by digital technologies. These digital methods are generating novel opportunities to collect individuals' permissions for use of private information but are blurring traditional boundaries of consent communication and documentation. Therefore, the rapid growth of digital technologies used for informed consent as well as the sheer volume of data resulting from electronic data capture are generating complex questions about individual engagement and data practices. This chapter presents emerging risks, benefits, and ethical principles about digital informed consent methods and technologies. For the areas where digital informed consent creates ethical uncertainties, ethical guidelines and user-design recommendations are provided.

INTRODUCTION

The practice of obtaining informed consent is a fundamental legal and ethical responsibility for organizations that provide personal services and for uses of personally identifiable information in many jurisdictions. The process of obtaining an individual's agreement starts with providing information for the individuals to make an autonomous informed decision about their options. This is followed by offering opportunities for individuals to ask questions, and culminates with obtaining their consent (De Sutter et al., 2020; Skelton et al., 2020). With advances in digital technologies, the informed consent process can be conducted electronically for the efficiency and convenience of pertinent stakeholders. In a 2019 eConsent survey on industry attitudes, more than 90% of respondents from biotechnology sponsors and

DOI: 10.4018/978-1-7998-8467-5.ch008

contract research organizations stated that they planned to implement electronic consent for some or all research in 2020 (Pundir et al., 2020). However, the emergence of digital methods has created opportunities to both enhance ethical protections and detract from intended protections.

While the nature of consent and permissions spans many industries, this chapter primarily focuses on informed consent processes that require a complex exchange of information, such as informed consent for healthcare delivery and participation in human subjects research. For the purposes of this chapter, health-related informed consent pertains to obtaining agreement for treatment, which may include medical, behavioral, and/or ancillary services (Anabo et al., 2019). Informed consent for research is an agreement between an individual and a researcher that includes specific elements of information, including procedures, risks, benefits of participation, and several more elements designed to assist an individual in making an informed decision (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979).

The authors' stance throughout this chapter aligns with Edenburg & Jones' (2019) "moral core" of consent: the adoption of emerging digital technologies for informed consent does not alter ethical principles, since these retain the moral compass shaped by societal values rooted in autonomy and justice. Digital informed consent methods instead require adaptations of consent processes and appropriate uses of permissioned data to adhere to ethical principles. While legal frameworks of consent vary across countries, this chapter limits discussions of privacy, individualism, human dignity, etc. to countries that have established governance to regulate and enforce these concepts (Edenberg & Jones, 2019). Last, for the purposes of this chapter, the terms "digital" and "electronic" informed consent are used interchangeably.

BACKGROUND

Ethical Principles of Informed Consent

There are many ethical principles from philosophy, medicine, and research that shape our values toward respecting decisions of other individuals and society. These person-centric approaches are intended to create an orderly society by establishing roles, responsibilities, standards, and expectations (Vallor, 2018). Several influential ethical codes were originally developed to deliver medical care and conduct human subjects research. Instrumental among them are the American Medical Association Code of Medical Ethics, the Principles of European Charter of Medical Ethics, The Belmont Report, and The Declaration of Helsinki, among others (American Medical Association, 2016; European Council of Medical Orders, 2011; The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979; World Medical Association, 2013). While there are some nuanced variations of each ethical guideline, common ethical elements include beneficence, nonmaleficence, justice, and respect for persons and their autonomy (Varkey, 2021).

"Beneficence" is a corollary concept to "nonmaleficence." Beneficence places a focus on the welfare of an individual first, while nonmaleficence is an obligation to do no harm (Varkey, 2021). They both promote ethical drive to put the individual's needs first. Such protections extend to help persons with disabilities and those who are isolated or at risk of harm (Varkey, 2021).

The principle of "justice" ascribes social obligations for fair and equitable distribution of risks and benefits among all populations (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). This principle promotes fair opportunities for access, bringing attention to the marginalization of certain groups and the risk that some groups may experience more loss of rights and privileges than others (Anabo et al., 2019).

An individual's "right to choose" is the fundamental belief underlying the principle of "respect for persons." Current practices of consent for treatment were derived from hard lessons of the past where the right to decide was withheld from individuals forced to undergo medical procedures involuntarily. Key to this principle is that the values of both parties are assumed to be aligned if the "autonomous" individual makes a choice and that choice is respected. Similarly, in conventional consent approaches, consent becomes an expression of individual empowerment (Schmietow, 2016). While the individual is central to the principle of autonomy, some decisions are also shaped by the individual's social context—a more nuanced understanding of individual decision-making incorporates interpersonal dynamics (Spruit et al., 2016). However, the "right to choose" was designed for individuals with reasoning capacity (Hallinan & Friedewald, 2015). Each person should not only be able to make their own choices according to one's values, but must also demonstrate the capability to make such decisions (McGraw et al., 2015). Additional protections are offered for individuals who lack decisional capacity (Varkey, 2021).

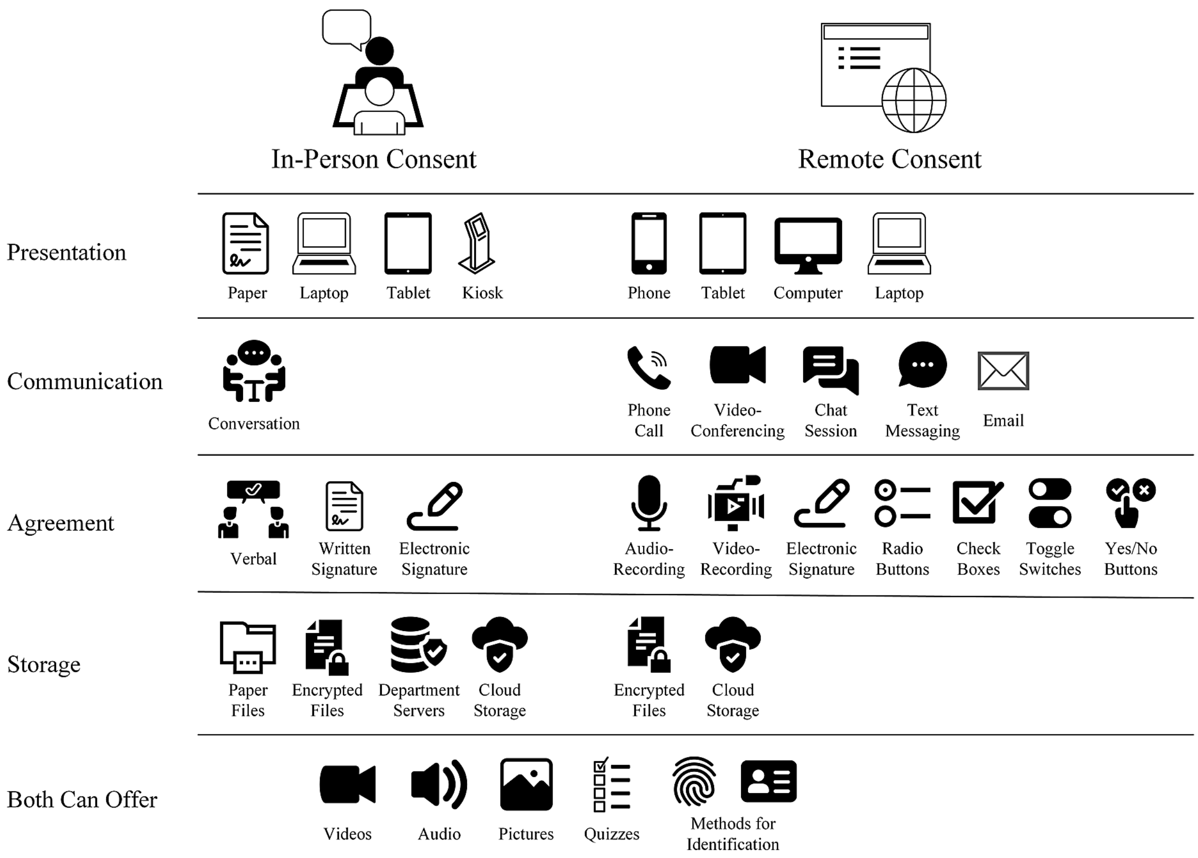
Autonomy can only be exercised when individuals have the opportunity to make decisions about themselves or their information when a consent process meets three elements: information, comprehension, and voluntariness (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). First, individuals must be presented with sufficient detail about their options to make informed decisions. "Information" should be presented in an unbiased manner that does not minimize risks or exaggerate possible benefits, and individuals should have the opportunity to ask questions prior to providing consent (Edenberg & Jones, 2019). These items typically include the purpose, nature of involvement, nature of data or specimens to be collected, anticipated risks and benefits, opportunity to ask questions, and option to withdraw at any time (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). Newer privacy requirements in the General Data Protection Regulation (GDPR) also specify that information presented for informed consent must be unambiguous and distinguishable from other requests (European Parliament and Council of the European Union, 2016c).

Next, "comprehension" of the content and context of consent language is believed to be as important as the information itself (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). This exchange has increasingly moved beyond signing an agreement (Anabo et al., 2019) and into a process where the end goal is to prepare the user for a greater understanding of information prior to making a decision (Litwin, 2016). Therefore, information should be presented in a manner consistent with the target population's language, maturity, and intelligence.

The last element of informed consent involves "voluntariness:" the ability to agree to participation and/or a set of choices without manipulation, coercion, or fear of penalty (Edenberg & Jones, 2019). To achieve voluntary agreement, individuals must have the opportunity to make a thoughtful, unhurried decision. Individuals should be given as much time as necessary to read information and to obtain a second opinion or consult with family and friends, if desired.

With the proliferation of large-scale data capture on the internet and with use of common electronic devices, there are new ethical questions regarding data privacy protections. While the concept of privacy was not originally included in major ethical codes cited earlier, a number of laws and ethical guidelines have been issued requiring explicit consent to ensure that an individual must first agree with the planned conditions for data processing (Tzanou, 2020). For the purposes of this chapter, data privacy is considered an expected outcome in an ethical informed consent process.

Figure 1. Uses of digital technologies to enable in-person and remote informed consent processes



Digital Technologies for Informed Consent

Nature of Technologies Used for Digital Informed Consent

A wide range of devices are used to present information with which to capture an individual's informed consent, depending on the nature of agreement and location of the respective parties. While the details of digital informed consent were originally presented with software downloaded on computers and laptops (Kakarlapudi & Mahmoud, 2021), informed consent is now increasingly presented and obtained on mobile devices. These devices primarily connect to internet-based sources to present information and capture responses for broader accessibility, while the touch-enabled interfaces are familiar and personalized for ease of use, making them ideal tools for interacting with consent information (Parsons & Abbott, 2013). Further, the prevalence and widespread use of smartphones enables ready access to provide consent (Parsons, 2015). Consent can be readily obtained in websites, software applications, and in electronic devices, including internet-connected "smart" devices and wearable technologies.

There are two primary circumstances driving the nature of technologies used for an informed consent process. As shown on the left side of Figure 1, digital technologies may be used to enhance the informed consent discussion during an in-person consent discussion. For example, tablet computers are provided to individuals in clinics to present a consent form to facilitate a consent discussion with researchers (Chhin et al., 2017), increase engagement with information (Jayasinghe et al., 2018), and manage storage of consent documents (De Sutter et al., 2020). When digital technologies are used to obtain informed consent in lieu of paper for an in-person exchange, digital technologies offer conveniences, but do not replace the interpersonal connection between individuals providing consent and the staff obtaining consent.

The second primary condition pertains to a remote exchange of information where it would not be practical to engage in an interpersonal interaction, such as when the consent process is conducted remotely. As shown on the right side of Figure 1, the digital informed consent process utilizes digital technologies for teleconferencing, videoconferencing, chat sessions, text, or email exchanges with organizations to facilitate communication (De Sutter et al., 2020). Most often, though, the digital informed consent process is designed to be conducted without direct interaction with the organization, such as web-based presentation of information to obtain agreement (Wilbanks, 2020). A low- or no-contact method of digital informed consent is common for obtaining agreement for internet-based cookies, collection of personal information from apps and/or use of other electronic technologies (Wilbanks, 2020).

Legal or Regulatory Requirements

As noted above, instrumental ethical codes in the United States (U.S.) and Europe have evolved over the past century to give individuals more protections and control for uses of personally identifiable information, including research and medical treatment (Edenberg & Jones, 2019). Newer legal requirements for organizations to obtain informed consent are driven by cultural developments toward privacy—especially in Westernized countries. European residents, for example, advocated for privacy as a "right," and any organization that wishes to collect and use personally identifiable information must first obtain explicit informed consent (European Parliament and Council of the European Union, 2016b). Germany advanced robust rights-based protections specific to digital consent (Federal Ministry of Justice and Consumer Protection, 2019). Several U.S. states have passed legislation inspired by GDPR to protect personal information, such the California Consumer Privacy Act (2018), that gives residents much more visibility into uses of their data and more informed choices about data uses. However, there is uncertainty about the extent of one's "right" to privacy. Clark et al. (2019) argue that privacy and the notion of informed consent should be balanced against benefits to society.

The process by which digital agreement is obtained is also regulated in the U.S. and Europe as referenced above to ensure consistent and appropriate protections. The characterization of digital technologies to capture an individual's agreement was first broadly defined in the Uniform Electronic Transactions Act (UETA, Uniform Law Commission, 1999) and the federal Electronic Signatures in Global and National Commerce Act (1999). Both Acts define an "electronic signature" as "an electronic sound, symbol, or process attached to or logically associated with a record and executed or adopted by a person with the intent to sign the record" (Uniform Law Commission, 1999, p. 5). The legislation was enacted to legitimize the validity of agreements made in electronic form. For example, UETA describes the legal validity of webpage click-through agreements. When individuals click "I agree," they perform an activity with the intent of demonstrating acceptance or agreement with specific terms.

For regulated research involving human participants, regulatory authorities in the United States have specified that informed consent may be provided using an electronic document. Specifically, an electronic consent form is permissible so long as the individuals or their legally authorized representatives are provided with a paper or electronic copy of the consent information (Office for Human Research Protections, 2016). However, for research regulated by the U.S. Food and Drug Administration, the system used to collect and process the electronic consent form must meet more stringent requirements for electronic records and electronic signatures (Food and Drug Administration & Department of Health and Human Services, 1997). Both vendors and organizations using an electronic system must meet administrative, procedural, and technical controls.

ETHICAL PRINCIPLES OF DIGITAL INFORMED CONSENT

As digital technologies are increasingly used to manage the process of informed consent, many factors complicate the ethical landscape of consent. The following section describes areas where digital technologies advance ethical principles of informed consent. After this section, the authors present factors that detract from the moral core of an informed decision. Both of these sections pertain to in-person and remote digital informed consent. Examples for in-person consent are drawn from medical or research contexts because in-person consent largely takes place in these settings, while examples for remote consent involve uses of the internet.

Areas Where Digital Informed Consent Advances Ethical Principles

This section provides an overview of individual-centric features that promote increased autonomy, comprehension, and engagement of information while promoting broader accessibility.

Autonomy

Among the most prominent ethical features of digital informed consent are the concepts of autonomy and the voluntariness of the consent process. Digital informed consent as an interactive media presents more opportunities to accept or decline participation (Royakkers et al., 2018). To make a truly informed decision, a digital informed consent process presents customizable relevant information with the freedom to select among a set of options. As examples, when visiting websites, individuals are presented with data usage options for website performance functions and marketing purposes. For research, individuals are presented with options to store data for varying durations and for different disease indications. These choices may be presented with toggle (on/off) switches, dropdown lists, checkboxes, and radio buttons or other features optimized for a user-friendly design (Anabo et al., 2019; Skelton et al., 2020).

Comprehension

Digital informed consent enhances the ethical concept of comprehension by allowing customization of information to be tailored to individuals' learning styles and preferences (Tait & Voepel-Lewis, 2015). For example, instead of using plain text to provide information, the digital consent process may include audio instructions where information is further explained (Parsons & Abbott, 2013). The recordings may be paused, replayed, or slowed down to promote comprehension.

Visual displays of digital informed consent offer the ability to include graphics, such as images, slideshows, and videos to enhance the comprehension of information (Parsons & Abbott, 2013). The images or videos allow the team or organization requesting consent to personalize the process and thus increase engagement with the team (Skelton et al., 2020). Short videos have been found to hold individuals' attention better than a text-based presentation (De Sutter et al., 2020) and facilitate better retention of content (Tait & Voepel-Lewis, 2015). These features are believed to reduce the prospect of "information overload" (Skelton et al., 2020), resulting in higher knowledge assessment scores than individuals presented with a consent form (Litwin, 2016). The presentation of information often also includes hyperlinks or buttons for individuals to press/select when seeking additional information about a topic (Skelton et al., 2020).

Engagement

A digital informed process promotes interactivity with written or verbal questions that increases engagement as well as test real-time understanding of the information (Tait & Voepel-Lewis, 2015). Consenting individuals can flag or highlight content to indicate additional questions or requests for clarification. These comprehension assessments provide explanations in response to incorrect answers and prevent progression through the consent process until the individual provides correct responses (Skelton et al., 2020). Further, the system can alert the organization or team of an individual's confusion of content material (Tait & Voepel-Lewis, 2015).

In contrast with one-time informed consent processes, digital consent platforms offer opportunities for ongoing interactions and communication between individuals and the organizations that obtain their informed consent. While an individual could contact an organization to withdraw or change their preferences for a paper consent form, the process requires considerable manual effort (Wee et al., 2013). Instead, a more ethical approach involves utilization of modern consent technologies. The term "dynamic consent" reflects that consent should involve ongoing autonomous choice where individuals dynamically change their preferences over time (Kaye et al., 2014). Using an online user interface, individuals select among a wider range of options (Schmietow, 2016) and modify their consent preferences whenever they change their minds (Wee et al., 2013). These options allow individuals to withdraw consent without the labor-intensive processes required for paper consent forms, thus providing individuals with greater control (Kaye et al., 2014).

The technology for ongoing and dynamic consent is also used for dynamic engagement to strengthen long-term treatment relationships with patients (Wee et al., 2013), interact with research participants for longitudinal studies (De Sutter et al., 2020), and provide results that might influence an individual's informed decision as to whether to maintain or revoke consent (Kakarlapudi & Mahmoud, 2021). Ultimately, digital dynamic consent options are believed to enhance trust in an organization (De Sutter et al., 2020).

Overall, a digital informed consent experience is more interesting than paper (Skelton et al., 2020) and increases the ethical principles of truly informed consent—better information, comprehension, and understanding of voluntariness (Skelton et al., 2020).

Inclusion of Underserved Populations

A digital process to obtain informed consent enables fair and equitable inclusion of some populations that would otherwise experience geographic or physical access barriers (Brall et al., 2019).

Specifically, online technologies bring opportunities to individuals, allowing for greater representation and increasing participation of underserved populations and/or hard to reach groups, such as rural communities (Clark et al., 2019). Individuals obtain remote healthcare, counseling, education, or other professional services by employing digital technology to enable informed consent discussions (Wilbanks, 2020). In underserved communities, expanded health services made possible by digital health exemplify the principle of justice (Brall et al., 2019).

Digital technologies offer an ethical informed consent process to accommodate a variety of physical and emotional disabilities. For example, digital platforms offer large text and multimedia presentation of information for individuals with limited vision (Kraft et al., 2019). Further, individuals unable to write due to injury or illness may touch their answer choices on a touch screen device and/or audiorecord their choices (Parsons, 2015). In addition, individuals with mental illnesses, such as severe social anxiety, provide informed consent virtually to receive a variety of services delivered electronically (Parsons, 2015).

Protecting the Integrity of Consent

When obtaining informed consent, an organization assumes the fundamental responsibility of protecting evidence of an individual's choices. For treatment, research, and privacy permissions, the consent agreement must be maintained for legal purposes (Parsons & Abbott, 2013). A digital informed consent process creates an audit trail of the consent status, promoting integrity and authenticity of the record (Kraft et al., 2019). The digital storage of the informed consent agreement also offers better version control and more effectively manages updates to the information (De Sutter et al., 2020). Further, the digital informed consent process provides easy access to the stored informed consent document (or information) for future reference, printing, or download with more reliable access than paper documents (Litwin, 2016).

Areas Where Digital Informed Consent Creates Ethical Uncertainties

While digital informed consent may promote many of the conveniences and efficiencies described above, there are circumstances for which digital technologies may detract from an ethical focus—especially in virtual settings. In this section, the authors question whether digital informed consent is truly informed and challenge the notion that when enhanced features are used, the digital technologies offer a better consenting process (Skelton et al., 2020). This section also highlights dilemmas between perceived convenience and ethics of a digital informed consent approach.

Is Digital Informed Consent Truly Informed?

Misconception of Comprehension

If an individual provided informed consent, it is argued that the person "must" have comprehended the information. However, ascertaining whether users comprehend the information is difficult and uncertain (Barrera et al., 2016; Clark et al., 2019). In clinical research, studies demonstrate that many participants lack understanding of key concepts (De Sutter et al., 2020), which may lead to a "therapeutic misconception" (e.g., thinking that a study procedure has direct therapeutic benefit) (Lidz et al., 2015). The structure and format of consent agreements may also create barriers for comprehension. These agreements are typically lengthy documents written in unfamiliar and vague language that people are unlikely to read thoroughly. Changing the format from paper to digital does not necessarily increase the likelihood that users will read these documents. Research on screen reading shows that people typically scan or skim the text and sign these agreements without having read them (Wilbanks, 2020). This widespread practice of accepting agreements without fully reading the terms may stem from having to keep up with rapidly changing websites technologies. These websites desensitize users to clicking buttons or ticking boxes without careful review of frequently modified terms of use (Barrera et al., 2016).

Literacy and Technology Barriers

The utilization of digital informed consent processes may lead to inequities that go beyond access to digital technologies and relate to individuals' limitations with technology proficiency (Brall et al., 2019). Depending on age, cognitive ability, health, and physical limitations, there are many individuals who are not able to navigate digital screens used for informed consent (Skelton et al., 2020). For example, some older adults have reported that they would need training and assistance to use a digitally-presented consent process, and others raised concerns about the physical demands of using hand-held devices (Jayasinghe et al., 2018). When considering these physical and technological limitations, there is risk of excluding certain individuals and populations from a digital informed consent process (Brall et al., 2019).

The transition from paper-based or in-person consent processes to a digital informed consent process creates inequalities when there is insufficient access to digital tools. While portable devices and cellular signals are becoming more prevalent (Pew Research Center, 2021), there are segments of the population excluded from using these technologies due to limited access to broadband internet or affordability of technological devices and services (Brall et al., 2019). Although people understand the benefits of digital technology adoption, they may not be in a position to access and use electronic consent options. Therefore, electronic consent alternatives should not serve as replacements for traditional informed consent methods (Simon et al., 2018).

Although digital technologies may offer efficient solutions to accommodate foreign languages (Mulder & Tudorica, 2019), physical and mental conditions/disabilities (De Sutter et al., 2020), and low literacy in education or health (Barrera et al., 2016) as described in previous sections, the implementation of such technologies is not straightforward. Litwin (2016) showed that non-English speaking patients were less likely to receive adequate information from a consent form—even in the presence of trained interpreters. Specifically, a digital approach that simply replaces in-person interpretation does not address this underlying problem. Furthermore, translating documents into other languages requires a quality assurance mechanism to ensure that terms and cultural concepts are represented appropriately (Mulder & Tudorica, 2019).

Missing Interpersonal Connections

While virtual informed consent has the opportunity to enable new features and outreach for the consent process, digital technology should not completely replace an interpersonal connection. Most informed consent processes involving complicated terms or procedures may still require a representative from the organization to answer questions and assess the individual's understanding (Kraft et al., 2019). Without the convenient ability for individuals to ask questions during the consent process, there is a legitimate ethical concern as to whether the individual can make a truly informed choice (Clark et al., 2019; Wilbanks, 2020).

Privacy Paradox

When individuals interact with information online, there is inconsistency between their stated preferences for privacy and their behaviors (Athey et al., 2017). Coined as the "privacy paradox," policy makers and ethicists have expressed concern that even when individuals are provided with internet-based choices for data uses and protections, they engage in behaviors that do not respect their opportunities for choice or privacy (Athey et al., 2017). Examples include accepting "terms of use" buttons or acceptance of web-based cookies without selecting among the information options presented. There is also risk that individuals may not play an active and autonomous role in what is intended to be an ongoing digital consent process. Consenting individuals instead perceive a burden of responsibility of maintaining data control and permissions in online settings (Vayena & Blasimme, 2017). Anabo et al. (2019) emphasize that the setting or context in which the information is presented influences individuals' expectations of privacy.

The requirement to allow data collection for product, website, or application use has led to an unethical power asymmetry between organizations and users who are not given any meaningful consent options for uses of their data (Vezyridis & Timmons, 2019). Specifically, this trade-off between using a desired service or product and the data requested in return may limit people's capacity to be completely "autonomous." Many informed consent terms with which people interact in their daily lives are non-negotiable, leaving individuals feeling resigned to whatever terms of service are offered (Edenberg & Jones, 2019). For example, information provided on websites creates a persistent digital footprint that is searchable and replicable (Anabo et al., 2019). Further, uses of mobile devices with cellular connections may inadvertently leak GPS location or collect data from other apps (Royakkers et al., 2018). Clark et al (2019) noted that a tweet shared on Twitter contains metadata about the user profile, device type, and the precise location where the tweet was posted. Individuals are also likely to be unaware that data from their uses of seemingly simple smart devices—televisions, thermostats, refrigerators, and even toothbrushes—are collected and sent to the manufacturer (Royakkers et al., 2018).

System Vulnerabilities

When collecting electronic data from individuals, organizations bear the ethical responsibility to secure and protect the data received. However, digital data processing creates risk of data vulnerabilities and breaches. GDPR describes a personal data breach as a "breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed" (European Parliament and Council of the European Union, 2016a). Nearly any electronic system can be hacked by exploiting software vulnerabilities, such as malware, or ploys on human behaviors, such as fraudulently assuming the identity of a user to gain unauthorized access (Royakkers et al., 2018). This risk is exacerbated by the fact that many mobile devices are connected to other devices. Royakkers et al. (2018) described an example where hacking into an internet-connected coffee machine provided access to the electronic door locking system. These types of breaches may be difficult to prevent as breaches could arise from anywhere in the world using the internet and modern technologies (Mulder & Tudorica, 2019).

Electronic data breaches may lead to a variety of social and economic harms (Anabo et al., 2019). Social harms may include inappropriate uses of data for discrimination, stigmatization, or loss of reputation (Brall et al., 2019). Economic harms could involve theft of funds or assets, including job loss or impact to one's insurability (De Sutter et al., 2020). These risks extend beyond the individuals directly affected and diminish trust within an entire community (Clark et al., 2019). Because of these security risks and resulting harms, some individuals are hesitant to use digital informed consent tools (Skelton et al., 2020).

Data Vulnerabilities

Uncertain Data Quality

Use of portable touchscreen devices for digital informed consent processes may also involve varying levels of data quality. Chalil Madathil et al. (2013) found that digital consent systems created more challenges for obtaining accurate data than paper-based systems. Similarly, individuals and organizational team members alike made more data entry errors using touchscreens than other formats (Chalil Madathil et al., 2013; Chhin et al., 2017). Data collection from the internet involves even more data quality challenges because individuals may misrepresent their identity or experiences without perceptions of accountability (Dike et al., 2019). Therefore, experts recommend employing quality control programs to assess data accuracy and authenticity if obtained from digital platforms.

Risk of Re-identification

Even when individuals provide informed consent for their electronic information to be collected and used in an "anonymous" manner, individuals may be unaware that their data could be retrieved and linked with other forms of personal information. For instance, researchers have determined that individuals are increasingly identified within aggregated and anonymized data sets ("Digital-Data Studies Need Consent," 2019). Data can be linked surreptitiously to other forms of personally identifiable information (Clark et al., 2019)—even when these anonymized data sets are incomplete (Rocher et al., 2019). There is additional risk for re-identification of individuals' genomic information, which leads to greater risk of identifying vulnerable populations such as members of racial or ethnic communities ("Digital-Data Studies Need Consent," 2019).

Methods to re-identify individuals are surprisingly simple. Methods such as data scraping, digital data matching, geolocation tracking, or data matching tools are achieved with minimal data such as simple demographics or Netflix subscriber movie ratings (Clark et al., 2019; Wilbanks, 2020). Hackers need little more than basic programming and analytics skills to link data sets (Narayan & Felten, 2014). Therefore, consent forms should not provide a "false promise" that data are de-identified (Chiauzzi & Wicks, 2019). Instead, individuals and organizations both should view privacy on a continuum of identifiability depending on the scope and nature of data.

Consequences for Failure to Follow Principles

While this section has focused on ethical concerns that arise when organizations do not obtain or honor terms of digital informed consent, these types of violations may also result in legal consequences—the majority of penalties thus far have been at the federal level. In the U.S., the Federal Trade Commission has brought over 70 cases against organizations that failed to obtain proper informed consent and/or misrepresented their practices for individuals' personal data (FTC, Federal Trade Commission, 2020). As an example, the FTC levied a $170 million judgement against YouTube and its parent company Google for collecting personal information about users watching programming directed toward children without first obtaining informed consent from their parents, a violation of the Children's Online Privacy Protection Act (Federal Trade Commission, 2020). As another example, the FTC issued a law enforcement action against Cambridge Analytica for falsely representing that an app would not collect identifiable information. Instead, the company had collected identifiable information from millions of Facebook users to perform voter profiling and targeted marketing (Federal Trade Commission, 2020).

SOLUTIONS AND RECOMMENDATIONS

With recognition that the emergence of digital consent demands new solutions, an integrative approach to digital ethical design and proactive behaviors is presented.

Ethical Framework

With the rapid growth of digital technologies and methods for online data collections, organizations must ensure ethical protections for individuals and their data. The nature of these protections reflects the organizations' values and the processes created for their business operations and compliance (Barrera et al., 2016). Ultimately, this framework should revolve around the "moral core" of protecting the rights and interests of the individuals they serve (Edenberg & Jones, 2019).

Such a framework can build on existing ethics codes, such as the ethics codes listed in the Introduction and the Association for Computing Machinery's Code of Ethics and Professional Conduct (Association for Computing Machinery, 2018). There are many excellent ethics frameworks available and free ethics toolkits are listed in the Additional Reading Section of this chapter. Such a framework should include several of the following ethical and risk mitigation strategies:

1. **Put individuals' welfare and best interests at the forefront**. This focus includes promoting individual autonomy and creating an environment of trust and respect.
2. **Recognize the unique ethical challenges in a digital technology landscape**.
3. **Create ethical responsibility and accountability** internally for organizations and externally to consenting individuals and the public.
4. **Identify the risks or conflicts inherent in your approach**. Vallor (2018) advocates a thorough "ethical risk sweeping" approach where a wide variety of organizational departments (e.g., legal, security, programming, finance, etc.) identify risks unique to their perspectives. Then the departments collectively create a comprehensive approach to manage risks throughout the entire software development life cycle. The author further emphasizes the need to expand the ethical circle to include perspectives outside of the typical group.
5. **Design for data privacy and security**. Sustainability of data privacy and security requires significant financial investment and organizational commitment.
6. **Document ethical standards and best practices**.
7. **Evaluate ethical approaches as an ongoing practice.**

These ethics approaches must go beyond regulatory compliance to demonstrate meaningful respect for individuals. The following sections go into detail about some of these principles.

Ethics by Design

Developers of digital informed consent technologies have an ethical responsibility to design technologies that respect individuals' rights and represent all parts of the population. Integration of ethical components in the planning stage is often referred to as "ethics by design" (Brall et al., 2019). The following subsections offer recommendations for presenting information, obtaining agreement, and designing the technology to promote a more ethical person-centered digital informed consent experience.

Honest and Transparent Communication

The conditions and terms of informed consent must be written in plain language in all contexts—without legal or technical jargon. The details should be presented in a concise and simple manner to encourage individuals to read the contents (Anabo et al., 2019). To achieve comprehension, contents should be offered in the language or languages with which the majority of readers would be most proficient (Brall et al., 2019). With recognition that there are varying levels of education among readers, the material should be written at a lower reading level, with many organizations striving for a 6th to 8th grade reading level for adults (Hadden et al., 2017). Both authors of this chapter worked at an academic medical center where patient-directed information was written at a 4th grade reading level to ensure comprehension.

To be lawful and fair, organizations should similarly demonstrate transparency regarding the information that could influence a person's willingness to provide informed consent (Mulder & Tudorica, 2019). Akin to requirements for informed consent content for treatment, human subjects research, and a variety of regulated disclosures, the content for digital informed consent should meet certain conditions to ensure adequate information is presented. These conditions include clear and honest communication for known activities, data collection uses, data sharing, and risks without burying certain terms to reduce recognition of less-than-desirable conditions (Mulder & Tudorica, 2019).

Nature of Data Collection and Use

Clark et al. (2019) advocate for not only explaining what information is collected but how information will be obtained—including hidden data collection, such as GPS tracking gleaning information from other apps—as well as why certain data are collected. Individuals should be told that information collected from commercial organizations may result in targeted marketing (Mulder & Tudorica, 2019) and that their data may be sold or used for other commercial purposes (Brall et al., 2019). Transparency of data use may also include concepts pertaining to which individual or organization will "own" (or serve as custodian of) the data, as well as where data will be stored and the length of time data will be stored (Brall et al., 2019). Information should also include how to withdraw access. To meet statutes such as the California Consumer Privacy Act (2018), organizations must be able to query and report use of specific individuals' data for a past length of time.

Nature and Scope of Risks

The information provided in a digital informed consent process should include details about the risks inherent in their choices and the likelihood of experiencing harm. The first component of this information may involve education about the possibility of harm from breaches or misuses of data. Barrera et al. (2016) found that one in five participants mistakenly did not realize there could be any risks from a digital consent process. As a result, the authors urge organizations to emphasize the nature of risks when there is potential for actual or theoretical harm. Some of these risks may pertain to emerging risks from processes not well known or established. Specifically, there are emerging risks from artificial intelligence algorithms increasing uses as part of data surveillance (Chiauzzi & Wicks, 2019; Clark et al., 2019). Chiauzzi et al. (2019) also advocate for transparency about the risk of re-identification so that individuals do not develop a false sense of security about "anonymous" data. The risk of re-identification is even greater when using genomic information. Even small sequences of genomic data may not only identify an individual, but also that person's genetically-related family members (Brall et al., 2019).

To identify potential risks, organizations should use ethics toolkits with checklists that advise about a wide range of potential risks. The Ethics and Algorithms Toolkit (Anderson et al., 2018), for instance, includes a worksheet that advises about risks to individuals' finances, emotions, reputations, safety, privacy, liberty, and intellectual property. To determine the best way to communicate higher risk categories, Chiauzzi et al. (2019) recommend consulting with members of the community—especially with high-risk groups—to determine methods for explaining and mitigating higher risks. Open communication about risk mitigation strategies is valuable to reduce individuals' concerns (De Sutter et al., 2020).

Communication Options

An organization is required to provide contact information for regulated research or privacy information (Mulder & Tudorica, 2019), but this is considered best practice for all digital informed consent processes. Traditional paper-based informed consent processes offer an email address or phone number, but digital informed consent processes offer more flexibility for contacting the organization requesting consent. Technology options include instant messaging services, online data entry contact fields, and even chatbots designed with artificial intelligence (Dike et al., 2019). Contact information that is easily accessible enables greater trust and assurance in the event they need help to address issues or complaints (Skelton et al., 2020).

Presentation Design Options

With digital presentation of information, user-centered design options include the ability to quickly adjust font type, size, and color as well as background color (Parsons & Abbott, 2013). Flexible presentations assist individuals with limited reading or language literacy, as well as those with poor eyesight (De Sutter et al., 2020). When designing a digital informed consent process that includes individuals with physical or literacy limitations, Brall et al. (2019) recommend including members of these populations to provide feedback during the design process to reduce the likelihood of inequities.

Further, the authors advocate for a flexible information design strategy that presents digital informed consent information with comfortable viewing options on smartphones, tablets, and computer screens (Kraft et al., 2019). User design features should include the ability to search for terms or keywords in the consent form (Skelton et al., 2020) and digitally highlight sentences or sections where there are questions so that these could be identified for interactive consultations with a person or a "virtual" AI-driven assistant (De Sutter et al., 2020).

A digital informed consent process should also be designed to include electronic mechanisms to assess an individual's understanding of the material. While this feature may not be needed for low risk agreements, an assessment promotes a more ethical approach when an individual may face more risk or commitment as part of the agreement. For optimal design function, comprehension should be assessed with quizzes with immediate feedback (Kraft et al., 2019). Further, the design should allow an individual to discuss options with family members, their physicians or attorneys prior to agreeing to the terms of the consent form (Skelton et al., 2020).

A digital consent document can also include features that do not allow individuals to agree to the terms unless the individual exhibits certain required behaviors. For example, to ensure the desired level of comprehension, individuals might not be permitted to progress to the next section of the consent form until they have achieved certain proficiency with quiz questions. They may also not be presented with agreement options until the person scrolls to the end of the text or spends a predetermined amount of time prior to being permitted to agree (De Sutter et al., 2020).

Consent Agreement Options

A digital consent process affords more flexibility with consent agreement options. Passive approaches such as pre-ticked checkboxes or opt-out features are no longer considered acceptable to meet newer privacy regulations that require evidence of more active intent (Brall et al., 2019; Rodriguez-Patarroyo et al., 2020). Active, intentional responses are believed to reinforce the ethical principle of voluntariness for agreement (Edenberg & Jones, 2019). Depending on the nature of consent process and risk to the individuals, active response options may include ticking checkboxes, toggle switches, typing one's name, signing one's name using a cursor, digitally signing a document with encryption protections, or verbally repeating phrases for a digital recording (Anabo et al., 2019; Rodriguez-Patarroyo et al., 2020).

It is also vital to design consent agreement options for individuals to change their choices or withdraw over time. Withdrawal or modification should be facilitated with a user-friendly interface with clear withdrawal features and explanations associated with all available choices (Anabo et al., 2019). Design features may necessitate that an identification code is associated with each individual's data to associate the request with the data to be removed (Anabo et al., 2019).

Security Solutions

Digital informed consent technologies should include stringent security protections and "privacy by design." These protections should include access restrictions and updated antivirus and antimalware protections (De Sutter et al., 2020) with secure end-to-end encryption of transmitted data (Skelton et al., 2020). Similarly, research has shown that individuals report greater trust in systems that utilize audit trails (Kakarlapudi & Mahmoud, 2021; Wee et al., 2013). Such audit trails should be designed to allow individuals to see how their data are being accessed and used. Further, the system should include features to protect individuals' identities. Such methods should capture the minimum necessary amount of personally identifiable information. To protect identifiable information, the system should use encryption, tokens, codes, or blockchain-based hashes to obfuscate identifiers and reduce the likelihood of identity theft (Kakarlapudi & Mahmoud, 2021).

Ethical Training

Because "ethics by design" principles are not necessarily taught to computer programmers and information technology staff (Barrera et al., 2016), these individuals would benefit from training about ethical thinking and approaches to individuals' needs and concerns about their data. Manufacturers designing software and organizations using the software alike could benefit from learning about social, cultural, and environmental factors that lead to individuals' understanding of information and design features that influence autonomy and voluntariness (Barrera et al., 2016).

FUTURE DIRECTIONS

Emerging Trends

While use of digital informed consent was already increasing in many industries over the past decade, the Covid-19 pandemic accelerated interest in transitioning medical and research informed consent processes toward remote and virtual technologies (Li et al., 2020). There was a relaxation of Health Insurance Portability and Accountability Act regulations and Centers for Medicare and Medicaid Services reimbursement guidelines to provide treatment using digital technologies when in-person visits could lead to unwarranted exposure (Robeznieks, 2020). The U.S. Centers for Disease Control reported that before the pandemic, 43% of healthcare facilities were capable of providing telehealth, but that percentage increased to 95% in 2020 (Demeke et al., 2021).

For research, there has been tremendous growth of digital and remote methods of informed consent due to patient vulnerability and social distancing (Goyal et al., 2021). In a survey of pharmaceutical and biotechnology executives conducted by Tufts University to determine the impact of Covid-19 on clinical research, the increased adoption of electronic informed consent was reported as the second largest emerging trend behind use of telehealth delivery (Le Breton et al., 2020). A pharmaceutical executive stated, "I expect eConsent and telemedicine will become routine and integral parts of clinical trials. They are pretty straightforward steps towards a more patient-centric trial, and could easily be adopted in most, if not all, protocols" (Le Breton et al., 2020, p. 6).

The continuing advances of digital technologies and the new opportunities to combine and utilize data require laws and regulations to be continually reviewed (Anabo et al., 2019). It has been a few years since stringent privacy regulations were implemented in Europe and some U.S. states, but there are ongoing struggles to understand and interpret the privacy regulations and the challenges with managing remote and web-based data collection. Of particular challenge is to maintain data protection when data leave the geographic boundaries, raising awareness and concerns of enforceability (Vayena & Blasimme, 2017). For research, the distribution of justice for uses of digital informed consent in international studies seems insufficient to reach distant communities (Anabo et al., 2019). The goal is to strike a balance between data protection and effective outreach.

Moving forward, there must not only be laws and regulations about digital informed consent but guidelines and interpretations. The current norms and practices have largely focused on security but require additional guidelines about justifiable processing and terms regarding when re-consent is required to go beyond the initial terms of agreement (Edenberg & Jones, 2019). These laws and regulations should be evaluated regarding their effectiveness at protecting individuals' rights and ethical interests (Edenberg & Jones, 2019). There are increasing calls for government bodies to promote standards and offer boundaries for ethical data uses and processing (Clark et al., 2019; Edenberg & Jones, 2019). Overall, the growing use of digital data requires ongoing education and a shared commitment toward addressing issues and cooperative approaches to protecting ethical protections (Clark et al., 2019).

Research

To expand the base of knowledge pertaining to ethical considerations for digital informed consent, research should be conducted in the following areas. Researchers utilizing digital technologies to present information via adaptable user interfaces should determine which digital features improve comprehension of the material and perceptions of voluntariness (Barrera et al., 2016). It is also important to understand individuals' ethical concerns and implications of collecting, storing, and sharing individuals' data. Research could elucidate how new levels of transparency and choice influence whether an individual would provide informed consent under different conditions (Clark et al., 2019). Additional studies could identify ethical risks and concerns that had not been previously considered for digital informed consent. These results could then inform organizations' ethical frameworks and provide direction for risk mitigation strategies.

CONCLUSION

As digital technologies have emerged that can empower individuals and create trustworthy interactions between individuals and organizations, there is need to establish sound ethical parameters for these digital informed consent processes. An ethical process must include all components of an informed decision-making process: information, comprehension, and voluntary participation by designing digital interfaces tailored to the nature of agreement and the population included (Skelton et al., 2020). An individual-centered process places the well-being and interests of the individual at the forefront and uses the technology to overcome a person's limitations to make a truly informed decision. Organizations designing digital informed consent methods must assume responsibility for obtaining agreement in an ethical manner. They can begin by creating ethical frameworks that delineate an organization's values, detail ethical assessments, and risk mitigation strategies that create an environment more likely to promote a trustworthy agreement and integrity with data protections and use. As digital technologies evolve, laws, regulations, and guidelines must change accordingly to maintain the individual at the center.

ACKNOWLEDGMENT

This book chapter received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The authors gratefully acknowledge the editing, comments, and suggestions from Chris Charles, Brooke Delgado, Audrey Goold, and Stacey McKenna, PhD that improved the quality of this manuscript.

REFERENCES

American Medical Association. (2016). *Code of medical ethics overview*. https://www.ama-assn.org/delivering-care/ethics/code-medical-ethics-overview

Anabo, I. F., Elexpuru-Albizuri, I., & Villardón-Gallego, L. (2019). Revisiting the Belmont Report's ethical principles in internet-mediated research: Perspectives from disciplinary associations in the social sciences. *Ethics and Information Technology*, *21*(2), 137–149. doi:10.100710676-018-9495-z

Anderson, D., Bonaguro, J., McKinney, M., Nicklin, A., & Wiseman, J. (2018). *Ethics & algorithms toolkit*. http://ethicstoolkit.ai/

Association for Computing Machinery. (2018, June 22). *ACM code of ethics and professional conduct*. https://www.acm.org/code-of-ethics

Athey, S., Catalini, C., & Tucker, C. E. (2017, September 27). *The digital privacy paradox: Small money, small costs, small talk*. National Bureau of Economic Research. https://www.google.com/books/edition/The_Digital_Privacy_Paradox/2AFfswEACAAJ?hl=en

Barrera, A. Z., Dunn, L. B., Nichols, A., Reardon, S., & Muñoz, R. F. (2016). Getting it "right:" Ensuring informed consent for an online clinical trial. *Journal of Empirical Research on Human Research Ethics; JERHRE*, *11*(4), 291–298. doi:10.1177/1556264616668974 PMID:27630213

Brall, C., Schröder-Bäck, P., & Maeckelberghe, E. (2019). Ethical aspects of digital health from a justice point of view. *European Journal of Public Health*, *29*(Supplement_3), 18–22. doi:10.1093/eurpub/ckz167 PMID:31738439

California Consumer Privacy Act of 2018. (2018). *Title 1.81., Sections 1798.100 - 1798.199.100*. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

Chalil Madathil, K., Koikkara, R., Obeid, J., Greenstein, J. S., Sanderson, I. C., Fryar, K., Moskowitz, J., & Gramopadhye, A. K. (2013). An investigation of the efficacy of electronic consenting interfaces of research permissions management system in a hospital setting. *International Journal of Medical Informatics*, *82*(9), 854–863. doi:10.1016/j.ijmedinf.2013.04.008 PMID:23757370

Chhin, V., Roussos, J., Michaelson, T., Bana, M., Bezjak, A., Foxcroft, S., Hamilton, J. L., & Liu, F.-F. (2017). Leveraging mobile technology to improve efficiency of the consent-to-treatment process. *JCO Clinical Cancer Informatics*, *1*(1), 1–8. doi:10.1200/CCI.17.00041 PMID:30657388

Chiauzzi, E., & Wicks, P. (2019). Digital trespass: Ethical and terms-of-use violations by researchers accessing data from an online patient community. *Journal of Medical Internet Research*, *21*(2), e11985. doi:10.2196/11985 PMID:30789346

Clark, K., Duckham, M., Guillemin, M., Hunter, A., McVernon, J., O'Keefe, C., Pitkin, C., Prawer, S., Sinnott, R., Warr, D., & Waycott, J. (2019). Advancing the ethical use of digital data in human research: Challenges and strategies to promote ethical practice. *Ethics and Information Technology*, *21*(1), 59–73. doi:10.100710676-018-9490-4

De Sutter, E., Zaçe, D., Boccia, S., Di Pietro, M. L., Geerts, D., Borry, P., & Huys, I. (2020). Implementation of electronic informed consent in biomedical research and stakeholders' perspectives: Systematic review. *Journal of Medical Internet Research*, *22*(10), e19129. doi:10.2196/19129 PMID:33030440

Demeke, H. B., Merali, S., Marks, S., Pao, L. Z., Romero, L., Sandhu, P., Clark, H., Clara, A., McDow, K. B., Tindall, E., Campbell, S., Bolton, J., Le, X., Skapik, J. L., Nwaise, I., Rose, M. A., Strona, F. V., Nelson, C., & Siza, C. (2021). Trends in use of telehealth among health centers during the COVID-19 pandemic— United States, June 26–November 6, 2020. *MMWR. Morbidity and Mortality Weekly Report*, *70*(7), 240–244. doi:10.15585/mmwr.mm7007a3 PMID:33600385

Digital-data studies need consent. (2019). *Nature, 572*(7767), 5–5.

Dike, C. C., Candilis, P., Kocsis, B., Sidhu, N., & Recupero, P. (2019). Ethical considerations regarding internet searches for patient information. *Psychiatric Services (Washington, D.C.)*, *70*(4), 324–328. doi:10.1176/appi.ps.201800495 PMID:30651058

Edenberg, E., & Jones, M. L. (2019). Analyzing the legal roots and moral core of digital consent. *New Media & Society*, *21*(8), 1804–1823. doi:10.1177/1461444819831321

Electronic Signatures in Global and National Commerce Act. (1999). *15 U.S.C. Ch. 96, Sections 7001-7031 of Public Law 106–229*. https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter96&edition=prelim

European Council of Medical Orders. (2011, June 10). *European charter of medical ethics*. http://www.ceom-ecmo.eu/en/view/principles-of-european-medical-ethics

European Parliament and Council of the European Union. (2016a, April 14). *GDPR, Article 4 - Definitions*. General Data Protection Regulation 2016/679. https://gdpr.eu/article-4-definitions/

European Parliament and Council of the European Union. (2016b, April 14). *GDPR, Article 7 - Conditions for consent*. General Data Protection Regulation 2016/679. https://gdpr.eu/article-7-how-to-get-consent-to-collect-personal-data/

European Parliament and Council of the European Union. (2016c, April 14). *GDPR, Recital 32 - Conditions for consent*. General Data Protection Regulation 2016/679. https://gdpr.eu/recital-32-conditions-for-consent/

Federal Ministry of Justice and Consumer Protection. (2019, November 20). *Federal Data Protection Act of 30 June 2017 (Federal Law Gazette I p. 2097), as last amended by Article 12 of the Act of 20 November 2019 (Federal Law Gazette I, p. 1626)*. https://www.gesetze-im-internet.de/englisch_bdsg/

Federal Trade Commission. (2020). *Privacy & security update for 2019*. https://www.ftc.gov/reports/privacy-data-security-update-2019

Food and Drug Administration & Department of Health and Human Services. (1997). Title 21 Code of Federal Regulations Part 11: Electronic Records; Electronic Signatures. *Federal Register*, *62*, 13464.

Goyal, M., Ospel, J. M., Ganesh, A., Marko, M., & Fisher, M. (2021). Rethinking consent for stroke trials in time-sensitive situations: Insights from the COVID-19 pandemic. *Stroke*, *52*(4), 1527–1531. doi:10.1161/STROKEAHA.120.031976 PMID:33588599

Hadden, K. B., Prince, L. Y., Moore, T. D., James, L. P., Holland, J. R., & Trudeau, C. R. (2017). Improving readability of informed consents for research at an academic medical institution. *Journal of Clinical and Translational Science*, *1*(6), 361–365. doi:10.1017/cts.2017.312 PMID:29707258

Hallinan, D., & Friedewald, M. (2015). Open consent, biobanking and data protection law: Can open consent be "informed" under the forthcoming data protection regulation? *Life Sciences, Society and Policy*, *11*(1), 1. Advance online publication. doi:10.118640504-014-0020-9 PMID:26085311

Jayasinghe, N., Moallem, B. I., Kakoullis, M., Ojie, M.-J., Sar-Graycar, L., Wyka, K., Reid, M. C., & Leonard, J. P. (2018). Establishing the feasibility of a tablet-based consent process with older adults: A mixed-methods study. *The Gerontologist*, *59*(1), 124–134. doi:10.1093/geront/gny045 PMID:29757375

Kakarlapudi, P. V., & Mahmoud, Q. H. (2021). A systematic review of blockchain for consent management. *Healthcare (Basel, Switzerland)*, *9*(2), 137. Advance online publication. doi:10.3390/healthcare9020137 PMID:33535465

Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2014). Dynamic consent: A patient interface for twenty-first century research networks. *European Journal of Human Genetics*, *23*(2), 141–146. doi:10.1038/ejhg.2014.71 PMID:24801761

Kraft, S. A., Garrison, N. A., & Wilfond, B. S. (2019). Understanding as an ethical aspiration in an era of digital technology-based communication: An analysis of informed consent functions. *The American Journal of Bioethics*, *19*(5), 34–36. doi:10.1080/15265161.2019.1587035 PMID:31090520

Le Breton, S., Lamberti, M. J., Dion, A., & Getz, K. A. (2020, October 22). *COVID-19 and Its impact on the future of clinical trial execution*. Applied Clinical Trials. https://www.appliedclinicaltrialsonline.com/view/covid-19-and-its-impact-on-the-future-of-clinical-trial-execution

Li, G., Yin, C., Zhou, Y., Wang, T., Chen, J., Liu, Y., Chen, T., Wang, H., Zhang, L., & Chen, X. (2020). Digitalized adaptation of oncology trials during and after COVID-19. *Cancer Cell*, *38*(2), 148–149. doi:10.1016/j.ccell.2020.06.018 PMID:32634378

Lidz, C. W., Albert, K., Appelbaum, P., Dunn, L. B., Overton, E., & Pivovarova, E. (2015). Why is therapeutic misconception so prevalent? *Cambridge Quarterly of Healthcare Ethics*, *24*(2), 231–241. doi:10.1017/S096318011400053X PMID:25719358

Litwin, J. (2016). Engagement shift: Informed consent in the digital era: Why electronic informed consent is key to supporting today's patient-centric mantra in clinical trials. *Applied Clinical Trials*, *25*(6-7). https://www.appliedclinicaltrialsonline.com/view/engagement-shift-informed-consent-digital-era-0

McGraw, D., Greene, S. M., Miner, C. S., Staman, K. L., Welch, M. J., & Rubel, A. (2015). Privacy and confidentiality in pragmatic clinical trials. *Clinical Trials*, *12*(5), 520–529. doi:10.1177/1740774515597677 PMID:26374682

Mulder, T., & Tudorica, M. (2019). Privacy policies, cross-border health data and the GDPR. *Information & Communications Technology Law*, *28*(3), 261–274. doi:10.1080/13600834.2019.1644068

Narayan, A., & Felten, E. W. (2014). *No silver bullet: De-identification still doesn't work*. Princeton University. https://www.cs.princeton.edu/~arvindn/publications/no-silver-bullet-de-identification.pdf

Office for Human Research Protections. (2016). *Informed consent FAQs*. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/informed-consent/index.html

Parsons, S. (2015). The potential of digital technologies for transforming informed consent practices with children and young people in social research. *Social Inclusion (Lisboa)*, *3*(6), 56–68. doi:10.17645i.v3i6.400

Parsons, S., & Abbott, C. (2013). Digital technologies for supporting the informed consent of children and young people in research: the potential for transforming current research ethics practice. *EPSRC Observatory for Responsible Innovation in ICT*. https://eprints.soton.ac.uk/id/eprint/356041

Pew Research Center. (2021, April 7). *Demographics of mobile device ownership and adoption in the United States*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Pundir, N., Lindroos, M., McDonnell, J., Byrom, B., & Egan, S. (2020). Delving into econsent: Industry survey reinforces patient centricity. *Clinical Researcher (Alexandria, Va.)*, *34*(1). https://acrpnet.org/2020/01/14/delving-into-econsent-industry-survey-reinforces-patient-centricity/

Robeznieks, A. (2020, March 19). *Key changes made to telehealth guidelines to boost COVID-19 care*. American Medical Association. https://www.ama-assn.org/practice-management/digital/key-changes-made-telehealth-guidelines-boost-covid-19-care

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, *10*(1), 3069. Advance online publication. doi:10.103841467-019-10933-3 PMID:31337762

Rodriguez-Patarroyo, M., Torres-Quintero, A., Vecino-Ortiz, A. I., Hallez, K., Franco-Rodriguez, A. N., Rueda Barrera, E. A., Puerto, S., Gibson, D. G., Labrique, A., Pariyo, G. W., & Ali, J. (2020). Informed consent for mobile phone health surveys in Colombia: A qualitative study. *Journal of Empirical Research on Human Research Ethics*. doi:10.1177/1556264620958606 PMID:32975157

Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, *20*(2), 127–142. doi:10.100710676-018-9452-x

Schmietow, B. (2016). Ethical dimensions of dynamic consent in data-intense biomedical research—Paradigm shift, or red herring? In D. Strech & M. Mertz (Eds.), *Ethics and Governance of Biomedical Research* (Vol. 4, pp. 197–209). Springer International Publishing. doi:10.1007/978-3-319-28731-7_15

Simon, C. M., Schartz, H. A., Rosenthal, G. E., Eisenstein, E. L., & Klein, D. W. (2018). Perspectives on electronic informed consent from patients underrepresented in research in the United States: A focus group study. *Journal of Empirical Research on Human Research Ethics; JERHRE*, *13*(4), 338–348. doi:10.1177/1556264618773883 PMID:29790410

Skelton, E., Drey, N., Rutherford, M., Ayers, S., & Malamateniou, C. (2020). Electronic consenting for conducting research remotely: A review of current practice and key recommendations for using e-consenting. *International Journal of Medical Informatics*, *143*, 104271. doi:10.1016/j.ijmedinf.2020.104271 PMID:32979650

Spruit, S. L., van de Poel, I., & Doorn, N. (2016). Informed consent in asymmetrical relationships: An investigation into relational factors that influence room for reflection. *NanoEthics*, *10*(2), 123–138. doi:10.100711569-016-0262-5 PMID:27478516

Tait, A. R., & Voepel-Lewis, T. (2015). Digital multimedia: A new approach for informed consent? *Journal of the American Medical Association*, *313*(5), 463–464. doi:10.1001/jama.2014.17122 PMID:25647199

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research* ((OS) 78-0013). U.S. Department of Health, Education, and Welfare. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf

Tzanou, M. (2020). *Health data privacy under the GDPR: Big data challenges and regulatory responses*. Routledge. doi:10.4324/9780429022241

Uniform Law Commission. (1999). *Uniform Electronic Transaction Act*. http://euro.ecom.cmu.edu/program/law/08-732/Transactions/ueta.pdf

Vallor, S. (2018). *Ethical toolkit*. https://www.scu.edu/ethics-in-technology-practice/ethical-toolkit/

Varkey, B. (2021). Principles of clinical ethics and their application to practice. *Medical Principles and Practice*, *30*(1), 17–28. PMID:32498071

Vayena, E., & Blasimme, A. (2017). Biomedical big data: New models of control over access, use and governance. *Journal of Bioethical Inquiry*, *14*(4), 501–513. doi:10.100711673-017-9809-6 PMID:28983835

Vezyridis, P., & Timmons, S. (2019). Resisting big data exploitations in public healthcare: Free riding or distributive justice? *Sociology of Health & Illness*, *41*(8), 1585–1599. doi:10.1111/1467-9566.12969 PMID:31423602

Wee, R., Henaghan, M., & Winship, I. (2013). Ethics: Dynamic consent in the digital age of biology: Online initiatives and regulatory considerations. *Journal of Primary Health Care*, *5*(4), 341–347. doi:10.1071/HC13341 PMID:24294625

Wilbanks, J. T. (2020). Electronic informed consent in mobile applications research. *The Journal of Law, Medicine & Ethics*, *48*(1, suppl), 147–153. doi:10.1177/1073110520917040 PMID:32342737

ADDITIONAL READING

American College of Healthcare Executives. (2021). Ethics toolkit. https://www.ache.org/about-ache/our-story/our-commitments/ethics/ache-code-of-ethics/creating-an-ethical-culture-within-the-healthcare-organization/ethics-toolkit

Anderson, D., Bonaguro, J., McKinney, M., Nicklin, A., & Wiseman, J. (2018). *Ethics & algorithms toolkit*. http://ethicstoolkit.ai/

Edenberg, E., & Jones, M. L. (2019). Analyzing the legal roots and moral core of digital consent. *New Media & Society*, *21*(8), 1804–1823. doi:10.1177/1461444819831321

Kraft, S. A., Garrison, N. A., & Wilfond, B. S. (2019). Understanding as an ethical aspiration in an era of digital technology-based communication: An analysis of informed consent functions [Review of Understanding as an ethical aspiration in an era of digital technology-based communication: An analysis of informed consent functions]. *The American Journal of Bioethics*, *19*(5), 34–36. doi:10.1080/15265161.2019.1587035 PMID:31090520

Kumra, R., Ranasinghe, E., Schlesinger, Y., & Wiggins, R. (2018). *Ethical OS: A guide to anticipating the future impact of today's technology.* https://ethicalos.org/wp-content/uploads/2018/08/Ethical-OS-Toolkit.pdf

McNamara, C. (2019). *Complete guide to ethics management: An ethics toolkit for managers.* https://managementhelp.org/businessethics/ethics-guide.htm

Mulder, T., & Tudorica, M. (2019). Privacy policies, cross-border health data and the GDPR. *Information & Communications Technology Law*, *28*(3), 261–274. doi:10.1080/13600834.2019.1644068

Skelton, E., Drey, N., Rutherford, M., Ayers, S., & Malamateniou, C. (2020). Electronic consenting for conducting research remotely: A review of current practice and key recommendations for using e-consenting. *International Journal of Medical Informatics*, *143*, 104271.

Vallor, S. (2018). *Ethical toolkit*. https://www.scu.edu/ethics-in-technology-practice/ethical-toolkit/

KEY TERMS AND DEFINITIONS

Autonomy: The ability to make an independent, self-governing decision based on an individual's assessment of information or situation.

Digital Data: Information created and stored in a computer mediated environment that can potentially be transmitted as discrete information signals over the internet, and may be subsequently processed and/or stored for a range of known and unforeseen purposes.

Digital / Electronic Informed Consent: Electronic systems which may incorporate multimedia in order to convey information and to obtain informed consent.

Dynamic Consent: The ability for individuals to independently change informed consent options over time.

Ethical Risks: Outcomes resulting in legal, cultural, economic, or reputational harm to individuals or organizations or that create moral controversies for other reasons.

Human Subjects Research: Research where a human subject is "a living individual about whom an investigator (whether professional or student) conducting research: 1) obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or 2) obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens." (45 CFR §46.102(e)(1))

Informed Consent: Informed consent is a fundamental ethical practice in biomedical research. It is the process of providing meaningful information to the potential participant in order to enable an autonomous well-informed decision on whether or not they wish to participate in the research study

Re-identification: The condition where data thought to be anonymous are linked with other data that allow individuals to be identified.

Section 4

Digital Ethics and New Realities

Chapter 9
Going Telemental:
Contact and Intimacy in Digital Mental Health

Shaun Respess
Virginia Tech, USA

ABSTRACT

Telemental health (TMH) is considered by many to be the future of mental healthcare, with some claiming that these methods should replace more traditional approaches. Early teletherapeutic initiatives demonstrate an immediate set of benefits for patients including improved access to care, reduced costs, better schedule flexibility, greater environmental familiarity, and higher rates of patient engagement. Notable limitations to TMH include enhanced privacy concerns, the variable digital literacy of certain populations/persons, and technological instability. However, other limitations regarding therapeutic relationships, experiences, and settings have gone undertheorized and are not sufficiently represented in the current research. This chapter surveys these considerations and argues that digital medical interventions are unable to effectively replicate the same degree of 'contact' and 'intimacy' available in physical care; providers should therefore be cautious in wholly replacing in-person methods or in implementing a standalone paradigm of digital care.

INTRODUCTION

As experts look to make treatment more accessible and efficient, mental health care facilities and services are undergoing an extensive digital revolution. Enduring issues of social precarity and inequality, along with new challenges presented by the COVID-19 pandemic, have forced practitioners to expand their use of digital technology to meet the diverse concerns of patients. Isolation and prolonged social distancing have since incited a "boom" in therapeutic videoconferencing, automated services, phone calls, texting, and social media (Kluger, 2020). These approaches constitute a paradigm of care known as *telemental health* (TMH) which offers unique techniques for persons to interact without needing to be physically present. As the need for these services continue, many experts are consequently considering

DOI: 10.4018/978-1-7998-8467-5.ch009

their long-term advantages and questioning their fit alongside of conventional therapeutic methods. In this chapter, the author applies elements from care ethics to explore and critique the sustainability of TMH, arguing that such services can compromise the quality of care even while providing several benefits. Furthermore, the author suggests that critical examinations of 'contact', 'intimacy', and embodied spaces should be crucial features of therapeutic assessment, including evaluations of TMH.

Care ethics is a normative theory that stresses the moral significance of connection, specifically the responsibilities that emerge from interdependent relationships. Of primary interest to this discussion are the ethical elements of care: attentiveness, responsibility, competence, and responsiveness (Tronto, 1993, 127). These features address different points of engagement in the caring/treatment process and expose significant limitations for TMH services. By focusing on connection, the elements allow us to consider the capability of persons to intimately share information, embody spaces together, and empathize with one another, points which are of great interest to many clinicians and counselors. They also help redefine health care ethics by challenging several assumptions of biomedicalization, most notably the depictions of illness and deficiency; the author understands persons in distress instead as functioning on a spectrum of disconnection. Care ethics provides a useful lens for carefully applying innovative services like TMH and for reflecting on the broader goals of mental health care. With respect to this conversation, digital methods must not be wholly dismissed nor provided free rein as their value is determined but should rather be given due caution and consideration.

THE LOGIC OF CARE

Mental health care is largely a product of biomedicalization, also colloquially known as the medical model. It is the "process by which more and more human problems and conditions have come to be defined and treated as medical problems and thus subject to medical study, diagnosis, prevention, treatment, and/or management... (bio)medicalization generally increases the power of biomedicine as an institution of social control." The pathologies and solutions of problems such as distress are also situated in individuals rather than in/from collective institutions (Gupta, 2019, 2; Conrad, 2007, 5-8). Ethical considerations are therefore a matter of patient choice and accountability, with practitioners obliged to avoid or minimize interference. This has been a rather promising direction for many healthcare specialties but difficult for the mental health field. Patients are simultaneously treated as self-governing *and* as incapable of acting in their own best interests due to their 'mental illness'. They are also regularly treated as sole decision-makers in situations where their personal relationships significantly influence their choices. Finally, they are managed according to undertheorized depictions of illness that sometimes foster fatalism; it is difficult to identify legitimate interests and solutions when one's condition is understood as inevitably restrictive.1

The medical model uses a "logic of choice," where the concept of patient choice acts "as a specific mode of organizing action and interaction; of understanding bodies, people and daily lives; of dealing with knowledge and technologies; of distinguishing between good and bad" (Mol, 2008, 7).2 It asserts that choice and equality are fundamental goods, meaning that patients should have fair opportunities to assess options for themselves. However, it offers no guidance for determining which treatments or solutions are preferable, instead leaving this responsibility to patients (74). This approach exposes patients as uninformed consumers and creates obstacles between patients and professionals who often struggle to interpret situations from the other's perspective. It also motivates providers to invest more in the marketing of biomedical products than in empirical deliberation (Veatch, 2008). Ethical principles like justice, autonomy, and non-maleficence are key contributions that have emerged from discussions of patient choice, but they are at present insufficient.3 As this chapter demonstrates, these deficiencies are only made more apparent by the challenges associated with TMH services. Instead, mental health should look towards a logic of care to motivate daily practice.

As Mol explains: "The logic of care is not preoccupied with our will, and with what we may opt for, but concentrates on what we do"; "the crucial moral act is not making value judgements, but engaging in practical activities" (7, 75). It alternatively observes "situations of choice" where the circumstances under which decisions are made can be uncertain, malleable, and challenging to overcome. Persons are understood not as self-sufficient consumers, but as interdependent selves situated in complex networks of relationships. Their ability to act and make choices stems from the past and ongoing care of others who have helped form and influence their well-being. Practices are evaluated by their specificity, adaptability, and collaboration, and are understood as open-ended processes without fixed timeframes or limits: "Care is not a transaction in which something is exchanged (a product against a price); but an interaction in which the action goes back and forth (in an ongoing process)" (Mol, 18). Mental health care treats diverse patients whose distress frequently affects and/or is caused by others around them. It can last for extensive periods of time depending on the needs of the patient and involves daily decisions that have lasting implications. The needs of patients are incredibly different and will manifest themselves in unique ways; their ability to make choices and their access to certain choices are also highly unequal. The logic of care is therefore better suited to address these challenges.

This explanation of care is largely descriptive, but the normative implications of care may be readily uncovered by reflecting on care practices. Past analyses by care ethicists have produced four concepts for consideration, what Tronto (1993) describes as the "ethical elements of care": attentiveness, responsibility, competence, and responsiveness. These correspond with different parts of the caring process: "caring about, noticing the need to care in the first place; taking care of, assuming responsibility for care; care-giving, the actual work of care that needs to be done; and care-receiving, the response of that which is cared for to the care" (127; Tronto, 2013, 22). Attentiveness means detailed attention to and acknowledgement of the needs of those dependent on us. Responsibility requires someone to assume accountability for ensuring that those needs are reasonably met. Those who do the work of caring must then do so with competence since, short of resource constraints, one may attempt to meet another's needs but do so ineffectively or inappropriately. Responsiveness appeals to the role of the cared-for in indicating their needs and reacting faithfully to the care they are receiving; their participation and feedback is a crucial part of whether a caregiver actually can be attentive and competent.4 These elements encourage experts to ask different questions and to identify suitable outcomes. They invite patients into discussions of illness (thereby promoting autonomy) and stress closer attention to their unique circumstances and situations. Principles of responsibility ensure that patients are neither ignored nor left behind. Lastly, competence motivates clinicians to emphasize empiricism when developing solutions and to communicate evidence-based treatments with more transparency and accuracy. These elements define the quality of care and thus serve as indispensable tools for assessing the efficacy of TMH.

Mental health care itself is a fluid science that must treat individuals as complex and indeterminate, despite its dependency on stable diagnostic categories that serve more as templates for the identification of symptoms than as necessarily accurate representations of mental distress. Professionals can observe patterns or abnormalities in one's 'expected' behavior and intervene in a myriad of ways; diagnoses and treatment are more so a matter of judgment than they are an exact science.5 Ethical considerations thus provide proper direction and boundaries, though they need not be purely restrictive. As Puig de la Bellacasa (2017) writes, "Constraints are not negative – enforcing – aspects of a practice; on the contrary, they are 'enabling' the practice, they make it specific, and develop in close relation to ways of being and doing" (152). Caring 'constraints' empower professionals to be more attentive, responsible, and competent, which in turn allow them to engage more fruitfully with highly diverse patients who might not "match well" with predetermined classifications. These constraints can then be used to assess the ability of professionals to develop sound observations, judgments, and solutions via digital technology.

As Barnes (2012) notes, attentiveness includes "recognition of the social and cultural circumstances and factors that affect the experience and nature of need" (20). Care is not just a "private concern" between individual persons but also a reflection of broader situations. Therefore, it is worth noting that here, distress is understood as a normative response to different forms of *disconnection*, defined by Hari (2018) as "being cut off from something we innately need but seem to have lost along the way" (82-83). Other scholars like Cvetkovich (2012) view conditions like 'depression' as public feelings: expressions of persons who "keep disappearing under the weight of daily life" (159).6 Types of disconnection include detachment from other persons, from meaningful work or activity, from a hopeful and secure future, from status and respect, from the natural world, and from life-enhancing values.7 People need purpose, security, community, activity, and assurance. Threats to these needs, especially within destructive sociocultural institutions, can compromise their psychological welfare. Sustainable treatments in this respect are difficult to facilitate since they require reconciliation with certain disruptive features that can be outside the control of treatment. Clinicians thus usually aim for symptom alleviation and short-term relief.8 Consequently, TMH services must be evaluated according to both immediate assistance and the long-term prevention of distress.

TELEMENTAL TRENDS

TMH is designed to "serve unmet health needs for professional resources, now aided by advancing capabilities of an ever-evolving and ubiquitous technology and the promise to improve access to quality healthcare while containing or restraining the rising cost of care" (Bashshur et al., 2016, 91). Essentially, TMH aims to bridge temporal, cultural, economic, geographical, and psychological gaps that prevent some persons from accessing care. These can include financial constraints, issues of stigma or shame with visiting counseling offices, and scheduling difficulties. Due to recent shutdowns related to COVID-19, they have also become the sole option for some. According to recent surveys, approximately 76% of mental health clinicians claim to only provide remote services at this time (American Psychological Association, 2020). They contend that TMH offers a low-cost, safe, and easy way to connect with existing clients while also allowing them to accept new persons experiencing distress.9 Providers enjoy greater access to clients and are able to meet more regularly than was formerly possible (Newsome, 2020). Born out of innovation and stimulated by necessity, TMH emerges as an attractive option for those seeking alternatives in a time of uncertainty and rampant disconnection.

'Telemental health' is used fairly interchangeably with *teletherapy*, which refers more generally to therapeutic counseling conducted via videoconferencing, phone calls, or texting. Despite the widespread equivalence, it is worth stating that there are some features of TMH that might not appropriately fit within conventional interpretations of therapy. For example, Abilify MyCite is a recently approved antidepressant used primarily for schizophrenia and bipolar disorder in addition to serving as a supplemental drug for some patients with depression. The drug contains a sensor that indicates whether it has been taken by connecting to a mobile application used by both patients and clinicians (Food and Drug Administration, 2017). Products like these are still appropriately contained within the class of TMH but would more aptly be categorized as medication management. One must also consider the range of popular meditation/mindfulness applications that patients can use to develop therapeutic strategies. These options involve less interpersonal interaction and instead provide users with tools to help heal themselves (Beard 2020).

The intended advantages of TMH services are well-documented and supported by the initial data available. Treatments like teletherapy produce lower service costs, offer greater schedule flexibility for both parties, have better environmental familiarity, report higher rates of patient engagement, reduce wait times and travel, and allow providers to reach more remote populations (Bashshur et al., 2016; Tutty et al., 2010; Langarizadeh et al., 2017; Zhou et al., 2020; Hilty et al., 2013). By eliminating travel and using exclusively digital appointment services, providers expand the range of options for their patients and themselves while eliminating a barrier for clients who may be less willing to schedule new sessions in-person or over the phone. TMH can also make care more affordable and mediate some of the financial constraints that prevent many from seeking assistance. This is surely a benefit to be commended in an era where mental healthcare is incredibly inequitable and difficult to access across the globe (Patel, 2012). In short, TMH has proven to increase convenience and help reach previously inaccessible and/or disadvantaged persons. Early reviews state that TMH appears to demonstrate a similar or equivalent ability to accurately diagnose patients and assess their symptoms and is comparable to in-person care for symptom alleviation. Certain endpoints like rehospitalization rates and the prevention of future symptoms will require more extensive research (Hilty et al., 2013, 451).

Most of the associated limitations of TMH are also apparent and can be anticipated given the digital modalities that make them possible. These services will likely exclude persons who are less proficient with platforms like *Zoom* or *Skype*, or who are prone to become frustrated when using social media or creating online appointments. Many will confront unreliable or inaccessible broadband and technological instability depending on the resources available to the client and clinician. In some studies, digital literacy and reliable internet service were the greatest obstacles to the expansion of telehealth (Langarizadeh, 2017, 244; Aboujaoude et al., 2015; Blandford et al., 2020). There are also plausible risks associated with the ableism and racism potentially embedded into the normative use of the technology: some are sensitive to or inhibited by the light and audio utilized by certain digital media, while others can have their differences in language or dialect distastefully amplified by phones or videoconferences. In an attempt to make care more accessible for some persons, TMH could exclude others who do not fit the primary archetype of these platforms.

Other limitations include fiscal responsibility and online privacy. While there are active efforts to implement and incentivize parity measures between TMH and in-person care, insurance companies control the terms by which they do or do not cover certain treatments and expenditures. The majority of insurers in the US (as well as in some other countries) apply location restrictions that can severely limit those who may access TMH services (Adams et al., 2018, 300). For those patients who are uninsured or underinsured, these treatments are actually *less* accessible. Issues of privacy plague TMH as well, particularly whether certain platforms are vulnerable to external interference. The security of social media, videoconferencing sites, and mobile networks owned by private companies is a frequent concern, one that will only be enhanced by the clinician's privileged access to a patient's personal information. Without an effective and independent set of regulations to protect health data, threats to privacy will likely be much higher among TMH services. It is worth mentioning that such privacy concerns are not limited to cyber-security. While many might enjoy a therapy session from the comfort of their own home, there are others who live in precarious environments that do not offer the same degree of ease. Clients who experience problems with persons they live with will have difficulty discussing those complications free from the inquisitive or accidental participation of those in their household. Furthermore, addressing these concerns could expose patients to additional and undue discomfort or harm. The author will elaborate on this problem later in the chapter.

The developing market of TMH will demand scrutiny as providers and clients surveil their options. Most will inevitably default to the options available or referred to them in accordance with their physical, economic, and cultural positions. Though the immediate effects of TMH appear mixed, this chapter appeals to the broader preferences of those receiving care: "Approximately 78% of rural patients and 72% of urban patients were 'moderately' or 'extremely' satisfied using TMH at the clinic. However, 44% of rural patients and just over 51% of urban patients strongly preferred face-to-face visits. Only 15.8% of rural respondents and 13.4% of urban respondents strongly preferred TMH visits" (Bashshur et al., 2016, 94). This conclusion is even more noteworthy: "participation rates also reflect that many adult participants (approximately one third) preferred face-to-face counseling. This finding may reflect the perceived value of in-person treatment features, such as eye contact, body posture, and touch" (Tutty et al., 2010, 234). It is thus crucial to examine the extent (if any) these conditions have on mental health treatment, including whether they may be sensibly accommodated by TMH.

CONTACT AND CONNECTION

Each element of care requires an appropriate amount of *contact* and connection between persons. According to Haraway (2013), the subjectivity of creatures is constituted within embodied relationships to others where their copresence, interactions, diverging associations of power, and intersecting customs or understandings meet. In her words, "meetings make us who and what we are in the avid contact zones that are the world. Once 'we' have met, we can never be 'the same' again" (287). Persons form and modify their complex personalities as well as meet many needs through these meetings, and there is some concern as to whether digital environments function as ideal 'zones' in this respect:

Most of us aren't getting a fraction of the person-to-person interaction we're accustomed to, and most of us are pretty well fed up with it. Virtual birthday parties are no party at all. Virtual happy hours have everything but the happy. Call it Zoom fatigue, cabin fever, flat-out loneliness—many today are suffering from isolation to one degree or another and long for the moment that the virtual lives we've been forced to live can be tossed aside. (Kluger, 2020)

Technology might help persons imitate or mitigate contact to some extent or in temporary bursts, but these benefits can occasionally fizzle out through extended exposure. Not only could the disconnection of clients be further exacerbated by digital media, but the platforms could stifle the engagement of both (or multiple) parties over time and prevent them from being sufficiently attentive and responsive to one another.

Through contact, Haraway claims that various knowledges are produced or amplified. These spaces of assemblage produce modes of "touch," though this analysis is careful to avoid necessarily conflating the term with direct physical contact in therapeutic settings. There are obvious and controversial problems with those who may abuse their position to facilitate inappropriate touching. Rather, touch refers to the embodied ability to inhabit a setting and physically interact with other creatures or things who also exist in that space. This ability initiates exchanges that are extremely informative and, for many, ward of disconnection. Touch creates informed affects: it in many ways accompanies and enables our capacity for empathy and sympathy: "In touch and regard, partners willingly are in the miscegenous mud that infuses our bodies with all that brought that contact into being... Caring means becoming subject to the unsettling obligation of curiosity; which requires knowing more at the end of the day than at the beginning" (Haraway, 36). Touch drastically influences the quality of caring exchanges by enabling persons to gather more information and engage more deeply with one another.

According to Puig de la Bellacasa (2017): "Involved knowledge is about being *touched* rather than observing from a distance... Touch therefore opens further meanings of knowledge that cares" (93). Touch is a "reaching out" that activates certain abilities and acquires information in ways that visual, audible, or other sensations cannot necessarily match. This is a prevalent concern given the known limitations of TMH methods, including an absence of visual cues with respect to telephones and audio services, a lack of audible cues in texting and social media, and notable time lapses over email and similar exchanges (Langarizadeh, 2017, 243). Individuals cannot quite "go back and forth" in the same manner as they could under physical settings, despite the technological advances and norms used to streamline communication. For many, digital interactions do not feel as authentic as embodied exchanges:

Virtual meetings lack many of the nuances that make in-person interactions feel connected and organic, while also presenting challenges such as internet connectivity issues, background noises, and awkward pauses or moments of cross-talk. As a result, those with many Zoom obligations may emotionally withdraw, becoming less participative in work meetings and choosing not to join video calls with friends despite already feeling socially isolated. (Sanderson et al., 2020, 260)

If patients cannot participate, stop participating, or participate with less responsiveness in these interactions, then carers will be unable to adequately fulfill the demands of attentiveness, responsibility, and competence. Though physical meetings cannot guarantee that patients will experience ideal conditions for participation, they normally encompass more meaningful sensations than can their teletherapeutic alternatives.

Caring relationships are generally *intimate*: "let us think of relations as intimate to the extent that interactions within them depend on particularized knowledge received, and attention provided by, at least one person – knowledge and attention that are not widely available to third parties" (Zelizer, 2009, 14). Good care involves a privileged and informed familiarity with those who receive care; in fact, patients need and normally expect special attention. Yet, imitating intimacy proves to be somewhat difficult from a virtual distance and can eliminate certain skillsets. These limitations can be harmful to intimate relationships where simple gestures, embodied contact, and empathic attentiveness make all the difference:

Just take the communication skills on which consultations depend: they are extensive. Pick the right words. Accept silences. Look at each other. Patients sit up straight or hunch their shoulders, a frightened or relieved look on their faces. Professionals smile, frown or search for something on their computer. Doctor and patient may lean together over the notebook with the results of blood sugar measurements. A nurse puts her hand on a patient's shoulder before she injects insulin. And then there are ever so many handshakes: consultations begin and end with one body touching another. Good communication *is a crucial precondition for good care. It also is care in and of itself. It improves people's daily lives. (Mol, 2008, 76)*

Without access to these skills, professionals are forced into precarious positions where their communication is diminished and where decisions are made without great confidence. TMH might allow clinicians to meet with previously inaccessible clients in remote locations and to meet with higher quantities of patients, but it appears to negatively influence their ability to maintain a certain quality of intimate connection with those clients. These issues are embedded in the technology and will require proper consideration before extensive use and recommendation of TMH.

SECURING SPACES

Most of TMH's restrictions are not necessarily indicative of poor execution by medical providers but are instead built into the digital infrastructures that define and discipline skills of engagement. Consider technological devices like touchscreens and trackpads or computer mice. Each "train us" to construct and relate to environments in distinct ways by placing constraints on what one may do and how they may do it. Similarly, platforms used in TMH orient users towards behaviors and knowledges reflective of the space's design and general preferences. Patients inhabit both physical and digital settings in a manner fitting to the media that facilitate interactions and do so differently as the modes of contact shift. Inhabiting a physical office will thus utilize different skillsets compared to those required by digital media. Elements of care must similarly acclimate to spaces as persons implement alternative skills and play by distinct rules of conduct; dispositions and actions are mediated by the limits placed on relationships. Consequently, TMH apparatuses may expedite the process of connection, but in doing so they also redefine its boundaries and expectations.

Environments not only facilitate care but require care themselves. The majority of counselors are all too aware of how essential it can be to create a relaxing, safe, and welcoming therapeutic space for their clients. Comfortable furniture, soft lighting, clean air, and general sanitation can mean quite a bit to patients. However, "there are many ways in which environmental design can include and exclude, and many ways in which lack of care for physical environments can contribute to a sense that these are risky spaces from which people may wish to exclude themselves" (Barnes, 2012, 135). Prime examples of 'risky' or exclusive spaces include those with physical obstacles that inhibit persons with disabilities from safely traversing and louder locations that trigger certain individuals with autism, post-traumatic stress disorder, or hyperacusis. In fact, these environments can be forceful stimulants for distress in their own right. Suitable spaces of care therefore must provide some alleviation. Given the constant concern that mental health professionals have for maintaining these spaces, TMH appears to provide a unique advantage in that digital platforms eliminate many precarious locations and relocate caring parties to spaces that they are likely to find more comfortable.

However, such a claim would readily dismiss digital platforms, applications, and media as concrete environments in and of themselves. As this chapter shows, this is surely not the case. In training individuals to be effective 'users' of a space, they also construct standards for preserving that space. Medical providers simply prefer digital apparatuses because they offload the responsibilities of preservation onto clients and/or private companies. This approach corresponds well with recent "self-care" initiatives where individuals are given a greater amount of responsibility for their well-being, all while public institutions and resources are steadily depleted (Ward, 2015).10 Digital platforms thus permit caregivers to disassociate from or eliminate a variety of embedded duties that are associated with their responsibility to patients. This is not an explicit goal of TMH, but rather a feature of the technological infrastructure which makes its methods possible. The constraints of digital zones are largely determined by private companies while the caregiving and cared-for parties essentially rent the space. “Caring for” in this sense is a more difficult endeavor since one has less control over the spatial conditions of care that could affect their patient.

Digital contact zones contain additional intermediaries that are more amenable to some populations while restrictive to others. Regardless, individuals shoulder the responsibility for managing their own digital literacy and accessing remote care. Those who are more vulnerable to technological instability, less familiar with the systems, or dissimilar from prototypical users are frequently “pushed out” of these zones and are discouraged from using them in the future. Issues of accessible broadband, cyber-security, and affordability are salient in these conversations. Deficiencies in embodied contact also mean that caregivers can miss cues from clients that could be of real significance. It may not be immediately obvious that a patient is not well-adjusted to the space and their frustration could cause tensions in the relationship or cause them to terminate treatment altogether. Therapeutic communication also suffers from common technical issues. For instance, “lag” can restrict one’s ability to convey distress or impede the counselor from properly understanding their expressions. Videoconferencing and other methods still lack shared peripheral cues; in an embodied space, objects or events that captures one’s attention are more easily noticed by others. These details might seem trivial, but they all can affect the competence and attention of users. Since the quality of care fundamentally depends on intimate knowledge, limitations to a clinician’s ability to observe physical cues, navigate shared environments, or respond in an appropriate manner can be especially restrictive.

TMH further redefines existing spaces like the home. For example, Oudshoorn (2012) argues that biomedicalization alters the behavior of others who coinhabit that home space: partners, siblings, children, and friends regularly take part in the daily examination and regulation of a patient’s body and mind. They might also feel encouraged to listen in on and participate in therapy. In some cases, this is a chance to invite new perspectives into the problem space. In other cases, cohabitants can threaten the emotional security of patients who are now unable to be as vulnerable around these abusive, negligent, or unsupportive others. It therefore is not ideal to encourage clients to divulge distress when precarious stimulants are nearby. The “privacy” of their intimate thoughts and emotions cannot be safely contained within the therapeutic relation and hence expose the patient to possible harm. Though practitioners can stay vigilant for these issues, digital spaces obscure them from understanding whether a foreign distraction is initiating discomfort, danger, or simply disinterest. This compromises the practitioner’s attentiveness and competence, fracturing their bond with patients. The absence of a mutually shared embodied space exposes limitations in TMH that are also particularly harmful in certain contexts. Though cyber-security is surely a risk worth critical attention, the emotional security of those involved in therapeutic relations must also be considered.

Without security, digital platforms and the physical settings we engage them from fail to adequately foster trust and solidarity whereas more conventional meetings might offer needed stability and relief. As Baier (1986) notes: “We inhabit a climate of trust as we inhabit an atmosphere and notice it as we notice air, only when it becomes scarce or polluted” (234). The fatigue, inauthenticity, and apathy that one may feel while using TMH could in part involve a disconnection from feelings of trust, stability, and hope, all of which are critical components of the therapeutic alliance. Finding a mental health professional who "fits" a client's needs and personality is difficult enough.11 Insecure spaces and sustained physical separation will stimulate these challenges and create new frustrations. They can deteriorate or annihilate whatever trust the participants might have had and thus lead to a dissolution of the therapeutic relation. It may be possible for practitioners to simply transmute the therapeutic space, but they must accept that certain cognitive, physical, and emotional benefits could be lost in the process. In terms of disconnection, digital spaces also perpetuate harm by reinforcing sentiments of isolation, frustration, embarrassment, and stress already plaguing those seeking treatment.

Finally, many patients report feeling restricted by the medicalization of their home and feel uneasy about their ability to "escape" their disorder/issue (Oudshoorn, 2012). TMH patients under these conditions may similarly feel trapped, thereby restricting their responsiveness and autonomy. A phenomenon known as biosociality, individuals tend to form particularly strong identities around their medical conditions and will relate to themselves accordingly (Rabinow, 2008). One is more likely to identify *as* a "schizophrenic," "depressed," or "bipolar" person rather than believe themselves to "suffer from" depression, schizophrenia, or bipolar tendencies. Their perceptions and actions tend to be more fatalistic, and they repeatedly refer to themselves as bystanders or victims to their neurological (or physical) activities. These are not inherently dangerous or wrong tendencies for one to have depending on the circumstances, but they can be detrimental to those who want more power over their distress or those who need sustainable relief. Biomedical spaces including the home can condition persons to interact with others, themselves, and their settings in ways that stimulate these tendencies. They could then foster more tentative patients who are less willing or able to be necessarily responsive. TMH approaches should consider not only who they might reach with digital spaces but also the types of persons they might create.

LESSONS AND REFLECTIONS

TMH offers flexible and innovative service options for both patients and providers to choose from but restricts the quality of care pertaining to embodied interpersonal benefits. They should be recommended as supplemental options in times of necessity or support but are likely not desirable as standalone forms of care for a wide range of patients. These methods have grown more favorable in response to recent concerns of cost-effectiveness, temporal efficiency, geographical accessibility, and consumer appeal, but prioritize these factors at the expense of crucial elements of care. The ability of a clinician to reasonably assess one's state, to take responsibility for their care, and to respond competently and with confidence is reduced, as is the ability of patients to be responsive and vulnerable. The biomedicalization of mental health care requires negotiations of this kind and TMH as a paradigm simply embraces its economic limitations more directly. However, as care it must be evaluated according to some minimal standard of use and efficiency. The author thus recommends elements of care as assessment tools and further suggests that emphasizing these elements exposes the significance of core features like contact, intimacy, and embodied spaces. TMH can be valuable for the future of mental health care if it is implemented with caution, purpose, and precision. To conclude then, this chapter will briefly discuss what discussions of TMH mean for mental health ethics and practice.

Early indications demonstrate that TMH is effective at reaching more remote populations and could accommodate those who are suddenly displaced or unable to participate in conventional meetings (such as those isolated during a pandemic). These are exceptional benefits that should prevent experts from merely disqualifying digital methods. Yet, in bridging these gaps the rollout of TMH services must prevent creating new exclusions based on technological literacy or environmental security. Committing to a comprehensive digital paradigm is not necessarily more inclusive, but rather entails an active choice in *who* is deserving of accessible or effective care. For these discussions, experts decide at a broader level who responsiveness and responsibility apply to by transitioning to a new subset of contact zones. Those who are not accustomed to or ready for these spaces will be (intentionally or unintentionally) left behind. In the meantime, there might be hybrid combinations of conventional and digital methods worth considering. Regardless, experts should critically evaluate the capability of digital techniques to *create* distance or omissions even as they rightfully meet the needs of those previously marginalized and/or excluded.

Second, conventional methods have much to gain from TMH when they are compared using these terms. For instance, one of the luxuries of TMH for many providers and clients has been participating from the comfortability of their own homes. Instead of simply moving care to the home, it is feasible to consider how they might bring 'home' to the therapeutic space. Gosselin (2020) makes a similar claim with respect to psychiatric hospitals: by developing "practices of home-making," hospitals reduce distress and increase compliance by treating these spaces as long-term comfortable environments where one may patiently heal. The use of "home-like" objects inspires positive cognitive associations, fosters familiarity, and reinforces trust. The home is also a special place where people can build meaning, ground themselves, and develop important networks of relations.12 Building on this idea, maybe professionals can use their interactions with remote clients to reimagine medical offices to be less awkward and more inviting. This is not a criticism of the many wonderful facilities who already diligently care for their spaces but rather an invitation to use TMH for additional reference points. Conventional offices could also reevaluate treatment affordability and schedule flexibility, as many barriers common to traditional methods may not only be inconvenient but also expensive and wasteful.

Lastly, TMH offers a unique opportunity to understand and navigate the fundamental features of spaces, media, and relationships in everyday mental health practices. These are underappreciated and vital conditions of care that radically affect its quality and meaning to persons. They may also be used to unpack the peculiar forms of disconnection that threaten the health, safety, and ambitions of patients seeking care, including those stimulated by certain treatment methods. Complacency and/or ambition can blind providers to the significance of concepts like contact and intimacy: specifically, how they are always present in care and how they might inform more effective, genuine, and equitable models of treatment. Care exists only in relationships where intimate information, opposing or collaborative values, and diverse needs are circulated in ongoing exchanges. Mental health care, including the innovations of TMH, is best served by embracing these facts even if it must accommodate several economic and social challenges. Whether digital technology will be an enabling or restrictive force moving forward is at present unclear, but it is certainly an intriguing source of controversy and exploration.

REFERENCES

Aboujaoude, E., Salame, W., & Naim, L. (2015). Telemental health: A status update. *World Psychiatry; Official Journal of the World Psychiatric Association (WPA)*, *14*(2), 223–230. doi:10.1002/wps.20218 PMID:26043340

Adams, S. M., Rice, M. J., Jones, S. L., Herzog, E., Mackenzie, L. J., & Oleck, L. G. (2018). TeleMental Health: Standards, Reimbursement, and Interstate Practice. *Journal of the American Psychiatric Nurses Association*, *24*(4), 295–305. doi:10.1177/1078390318763963 PMID:29589800

American Psychological Association. (2020, June 5). *Psychologists Embrace Telehealth to Prevent the Spread of COVID-19*. https://www.apaservices.org/practice/legal/technology/psychologists-embrace-telehealth

Baier, A. (1986). Trust and Antitrust. *Ethics*, *96*(2), 231–260. doi:10.1086/292745

Barker, P. (2011). *Mental Health Ethics: The human context*. Routledge.

Barnes, M. (2012). *Care in Everyday Life: An Ethic of Care in Practice*. Policy Press.

Bashshur, R. L., Howell, J. D., Krupinski, E. A., Harms, K. M., Bashshur, N., & Doarn, C. R. (2016). The Empirical Foundations of Telemedicine Interventions in Primary Care. *Telemedicine Journal and e-Health*, *22*(5), 342–375. doi:10.1089/tmj.2016.0045 PMID:27128779

Beard, C. (2020, October 1). *Peace of Mind: There's an App for That.* McLean Hospital. https://www.mcleanhospital.org/essential/peace-mind-theres-app

Beauchamp, T. L., & Childress, J. F. (2001). *Principles of Biomedical Ethics* (5th ed.). Oxford University Press.

Blandford, A., Wesson, J., Amalberti, R., AlHazme, R., & Allwihan, R. (2020). Opportunities and challenges for telehealth within, and beyond, a pandemic. *The Lancet. Global Health*, *8*(11), e1364–e1365. doi:10.1016/S2214-109X(20)30362-4 PMID:32791119

Conrad, P. (2007). *The Medicalization of Society: On the Transformation of Human Conditions into Treatable Disorders*. JHU Press.

Cvetkovich, A. (2012). *Depression: A Public Feeling*. Duke University Press.

Engster, D. (2007). *The Heart of Justice: Care Ethics and Political Theory*. Oxford University Press.

Food and Drug Administration. (2017, November 13). *FDA approves pill with sensor that digitally tracks if patients have ingested their medication.* FDA News Release. https://www.fda.gov/news-events/press-announcements/fda-approves-pill-sensor-digitally-tracks-if-patients-have-ingested-their-medication

Gosselin, A. (2020). At Home in a Psychiatric Hospital. *Social Philosophy Today*, *36*(July), 71–87. doi:10.5840ocphiltoday202012971

Gupta, K. (2019). *Medical Entanglements: Rethinking Feminist Debates about Healthcare*. Rutgers University Press.

Haraway, D. J. (2013). *When Species Meet*. University of Minnesota Press.

Hari, J. (2018). *Lost Connections: Uncovering the Real Causes of Depression and the Unexpected Solutions*. Bloomsbury Publishing USA.

Hilty, D. M., Ferrer, D. C., Parish, M. B., Johnston, B., Callahan, E. J., & Yellowlees, P. M. (2013). The Effectiveness of Telemental Health: A 2013 Review. *Telemedicine Journal and E-Health, 19*(6), 444–454.

Holm, S. (2019). Bioethics and mental health—An uneasy relationship. *Ethics. Medicine and Public Health*, *10*, 1–7.

Kluger, J. (2020, August 27). Online Therapy, Booming During the Coronavirus Pandemic, May Be Here to Stay. *Time*. https://time.com/5883704/teletherapy-coronavirus/

Langarizadeh, M., Tabatabaei, M. S., Tavakol, K., Naghipour, M., Rostami, A., & Moghbeli, F. (2017). Telemental Health Care, an Effective Alternative to Conventional Mental Care: A Systematic Review. *Acta Informatica Medica*, *25*(4), 240–246. doi:10.5455/aim.2017.25.240-246 PMID:29284913

Mol, A. (2008). *The Logic of Care: Health and the Problem of Patient Choice*. Routledge. doi:10.4324/9780203927076

Newsome, M. (2020, June 3). *Teletherapy in the Age of COVID-19*. North Carolina Health News. https://www.northcarolinahealthnews.org/2020/06/03/teletherapy-in-the-age-of-covid-19/

Oudshoorn, N. (2012). How Places Matter: Telecare Technologies and the Changing Spatial Dimensions of Healthcare. *Social Studies of Science*, *42*(1), 121–142. doi:10.1177/0306312711431817 PMID:22530385

Patel, V. (2012). Global Mental Health: From Science to Action. *Harvard Review of Psychiatry*, *20*(1), 6–12. doi:10.3109/10673229.2012.649108 PMID:22335178

Puig de la Bellacasa, M. (2017). *Matters of Care: Speculative Ethics in More than Human Worlds*. University of Minnesota Press.

Rabinow, P. (2008). Artificiality and Enlightenment: From Sociobiology to Biosociality. In J. X. Inda (Ed.), *Anthropologies of Modernity: Foucault, Governmentality, and Life Politics* (pp. 179–193). Blackwell Publishing.

Sanderson, W. C., Arunagiri, V., Funk, A. P., Ginsburg, K. L., Krychiw, J. K., Limowski, A. R., Olesnycky, O. S., & Stout, Z. (2020). The Nature and Treatment of Pandemic-Related Psychological Distress. *Journal of Contemporary Psychotherapy*, *50*(4), 251–263. doi:10.100710879-020-09463-7 PMID:32836377

Tronto, J. C. (2013). *Caring Democracy: Markets, Equality, and Justice*. NYU Press.

Tronto, J. (1993). *Moral Boundaries: A Political Argument for an Ethic of Care*. Routledge.

Tutty, S., Spangler, D. L., Poppleton, L. E., Ludman, E. J., & Simon, G. E. (2010). Evaluating the effectiveness of cognitive-behavioral teletherapy in depressed adults. *Behavior Therapy*, *41*(2), 229–236. doi:10.1016/j.beth.2009.03.002 PMID:20412887

Veatch, R. (2008). *Patient, Heal Thyself: How the "New Medicine" Puts the Patient in Charge*. Oxford University Press.

Ward, L. (2015). Caring for Ourselves? Self-Care and Neoliberalism. In M. Barnes, T. Brannelly, L. Ward, & N. Ward (Eds.), *Ethics of Care: Critical Advances in International Perspective*. Policy Press. doi:10.2307/j.ctt1t89d95.8

Whooley, O. (2017). Defining Mental Disorders: Sociological Investigations into the Classification of Mental Disorders. In T. L. Scheid & E. R. Wright (Eds.), *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems* (pp. 45–65). Cambridge University Press. doi:10.1017/9781316471289.006

Woo, B., Walton, E., & Takeuchi, D. T. (2017). Cultural Diversity and Mental Health Treatment. In T. L. Scheid & E. R. Wright (Eds.), *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems* (pp. 493–511). Cambridge University Press. doi:10.1017/9781316471289.029

Zelizer, V. A. (2009). *The Purchase of Intimacy*. Princeton University Press. doi:10.1515/9781400826759

Zhou, X., Snoswell, C. L., Harding, L. E., Bambling, M., Edirippulige, S., Bai, X., & Smith, A. C. (2020). The Role of Telehealth in Reducing the Mental Health Burden from COVID-19. *Telemedicine Journal and e-Health*, *26*(4), 377–379. doi:10.1089/tmj.2020.0068 PMID:32202977

ENDNOTES

1 Holm (2019) unpacks these challenges in greater detail, suggesting that the field of bioethics has much to learn from complications in mental health practice. These revelations are salient given the vast numbers of people who suffer from chronic mental disorders. As Barker (2011) further notes: “ethics is only meaningful where people – or groups of people – are self-governing and have the opportunity to make choices free from any coercion. Rarely is this the case in the mental health field. The limits imposed on a person’s exercise of freedom – however explicit – continue to haunt contemporary practice” (3).

2 Mol (2008) uses the term “logic” rather loosely. The approach is not categorically coherent or fixed but essentially states that: “Events somehow tend to fit together, there are affinities between them”; logic is understood by her as the rationale behind “modes of ordering” (8).

3 Healthcare ethics are largely informed by Beauchamp and Childress’s (2001) principles of medical ethics: autonomy, non-maleficence, beneficence, and justice. These are supported by four rules: veracity, privacy, confidentiality, and fidelity.

4 Engster (2007) notes that attentiveness means asking, “Do you need something?” while responsiveness then directs us to ask, “What do you need?” (30-31).

5 Classifications of mental disorders cannot be reliably attributed to biomedically discerned attributes but are rather the product of expert consensus in how symptoms and ‘diseases’ should be understood (Whooley, 2017, 47).

6 Cvetkovich (2012) continues to assert that we are at an impasse with what clinical research can reasonably show us: “We don’t need scientific research to explain what’s going on; we need better ways of talking about ordinary life, including the dull feelings of just getting by” (159).

7 Not included in this description are genetic predispositions that cannot *cause* distress but may make certain synapse assemblages more likely. One’s neurological changes are more accurately understood as synaptic conditioning that responds to exterior stimuli. Prolonged exposure to discomfort or pain may train the brain into proactively activating those synapses which correspond with emotions such as fear or anxiety (Hari, 2018, 146-148).

8 Despite the conflict between sustainable healing and immediate relief, both goals are extremely important. Gupta (2019) identifies this problem within mainstream medical interventions that simultaneously alleviate some individual suffering while enhancing social inequality. In her analysis, she claims that relief and survival frequently come from paradigms of normalization, including those induced by treatment.

[9] Zhou et al. (2020) specifically argue that TMH services are "perfectly suited to this pandemic situation – giving people in remote locations access to important services without increasing risk of infection" (378). They, along with Sanderson et al. (2020), report a surge in persons seeking mental health treatment as a direct result of pandemic-related distress and disruption.

[10] Ward (2015) argues that 'self-care' is an initiative that neoliberal policy makers use to reduce support for public welfare and increase responsibility for individual citizens. Rather than subsidize or enhance public modes of healthcare, for instance, individual citizens are forced to "stay healthy" and avoid harm.

[11] Within therapy, a patient's fit with their therapist or their "cognitive match" is a crucial part of the therapeutic alliance. This can include not only ethnic and cultural similarities, but also alignment in beliefs, perspectives, and experiences (Woo et al., 2017, 503-505).

[12] Gosselin (2020) covers these home-making practices and objects in greater detail. Since one effectively "moves into" a psychiatric hospital for a temporary or extended period of time, some of these suggestions will reasonably not apply to other therapeutic spaces.

Chapter 10
The Impact of Decentralized Technologies on Social Media Megacorporations

Richard Foster-Fletcher
MKAI, UK

Odilia Coi
MKAI, UK

ABSTRACT

Social media is a mega-industry built by systematically monetizing the exploitation of human emotions, reactions, and biases. The authors explain how this industry became so profitable by creating a fear of missing out (FOMO) to command our attention, blending news and content in one feed to keep users 'in-app', and using powerful algorithms to promote more provocative posts, filter content, and trigger the reward centres of our brains. The authors examine how decentralized technologies, including cryptocurrencies, tokenization, and blockchain are being developed and deployed into new social media applications. The authors speculate on how these blockchain-backed startups could challenge the status quo and appeal to new expectations of user privacy, tighter regulation, and a more equitable monetization system.

INTRODUCTION

The most successful social media companies have extraordinary abilities to extend usage and extract user data without causing alarm. They have profited from what has been called surveillance capitalism. Successful social media companies do not charge users for access as they make their money downstream from providing targeted advertising.

The plain truth is that it is not feasible for a Social Media business to make significant revenues without harvesting user data, predicting preferences and behavior, and using these insights to sell target advertisements. Just as Google is not merely a search engine; it is a digital concierge service attempting

DOI: 10.4018/978-1-7998-8467-5.ch010

to anticipate our requests or interests; Facebook is not a community tool to safely share information and opinions. It is a phenomenally successful tool for monetizing the systematic exploitation of human emotions, reactions, and biases. Social Media is most successful commercially when the information that we, the users, see validates our beliefs and identities. This knowledge has encouraged the platform owners to set their algorithms to promote content that we are most likely to find interesting. This algorithmic practice can create digital content echo chambers that reinforce our opinions and polarize our views. The more we engage with social content, the more personal data we provide to the platform owners. The more data the platform has, the better it can predict our behaviors, intent, and interests. The better these predictions are, the better the value proposition is to advertisers. Facebook has mastered this process; it has grown revenues from $7.87bn in 2013 to $86.7bn in 2020, with 98% of the revenue coming from advertisers. The internet is monopolized by a tiny number of massive corporations. Internet search is essentially wholly controlled by Google, with 90% of search queries going through their data centers. Google's web browser Chrome had a 0.52% market share in 2008, but by 2020 66.59% of desktop and 61.35% of mobile users used it as their default web browser.

Facebook generates 4 petabytes of data per day. Every 60 seconds, 510,000 comments are posted, 293,000 statuses are updated, 4 million posts are liked, and 136,000 photos are uploaded. Google processes over 20 petabytes of data per day, recording every search and every YouTube video watched. Google Maps logs every route and journey.

Google has a centralized data-management strategy, the default strategy for large technology companies. Cloud providers opened 15 million sq. ft of data center space in 2020, and Facebook operates 18 massive data centers in 17 regions. Massive amounts of data require massive amounts of central storage.

History has taught us that too much accountability in a single position of power almost always widens the threat of abuse. As Ben Dickson writes in VentureBeat:

If the servers of these entities go down, we lose access to vital functionality. If they get hacked, we lose our data. If they decide to monetize our data in unlawful ways or hand it over to government agencies, we likely won't learn about it. If they decide to censor or prioritize content based on their interests, we won't be able to do anything about it (Dickson, 2017, para. 4).

Governments around the world are aware of Big Tech data monopolies; they were slow to wake up to this, but once they realized how powerful these companies were becoming, they took steps to try and reduce the control and power the private companies have over user data. Over this decade, global governments will form coalitions with the intent of reclaiming power. The critical role of Governments is to force companies and markets to account for the impact they have on society. Legislations force innovation, and if Governments can successfully force these companies to account for societal impact, they could innovate differently by prioritizing privacy, security, and profit. Technology companies, however, have vast cash reserves and can soak up fines and legal costs for many years to come before we see industry-wide changes.

We may not have to wait, however, for external governance. In the much nearer term, it could be the impact of decentralized technologies powered by blockchain that change the Social Media landscape. Centralized data centers are prone to infiltration, and Web 3.0 is likely to be a decentralized web. A new version of the web that spreads the power load, data, and decision-making across many independent machines.

Dickson (2017) goes on to describe this possibility:

In a fully decentralized internet, instead of one or a few organizations running the system, a community of users and a network of independent machines would own and power these vital services. This would make them more resilient to failures and hacks while ensuring no single entity can use them in nefarious ways (para. 6).

A new decentralized web would employ a transparent and community-driven decision-making process to align all parties' interests. It will provide a level of security and transparency that the Social Media industry must embrace if it loses consumer trust. Decentralization is not currently desirable for the largest technology companies. Each company's centrally held data is its most valuable asset; Google and Facebook's cannot decentralize without remaking their businesses. As Kodak, Xerox and Blockbuster demonstrated, remaking a business is hard.

BACKGROUND

To establish how decentralized technologies could topple social media empires, we must first look back at how and why specific social platforms could scale to billions of users. This context will allow us to see more clearly how new market forces will disrupt the current mega-corporations this decade.

Approximately ten years after the first electronic mail (email), several private companies noticed an unmet need for allowing internet users to establish a personal 'digital home' and broadcast information to their friends. The internet transitioned from version 1.0 to version 2.0, a user-orientated web that created the stage for the first Social Media platforms. Websites where users create an information-rich profile, share this with a list of contacts, and manage many interpersonal relationships. These social apps began as connection-based tools like the short-lived Sixdegrees.com and Friends Reunited, which allowed users to find friends and share text, photos, and messages.

At the turn of the millennium, there were other early contenders in the industry, but the two most dominant platforms were Friendster (founded in 2001) and MySpace (launched in 2003). Friendster briefly looked like it would emerge as the social networking application winner and was so confident about its success that it turned down a $30m offer from Google. As it chased more and more users, however, the number of bugs in the system started to make it unusable, and eventually, out of frustration, the 10 million users migrated elsewhere.

Myspace was briefly the undisputed king of social media, but it faced user trust problems due to the aggressive advertising campaigns on the user's profile. Under pressure to deliver a profit, MySpace executives compromised the user experience with intrusive and sometimes questionable advertising (Garrahan, 2009). One of the most famous advertisements was 'Punch The Monkey' that invited users to click on it to win a prize. Instead of providing a reward, however, the app redirected to a survey or a request to enter credit card information.

Each early social network shone brightly for a while before its flame extinguished. hi5 (founded in 2003) and Orkut (created in 2004) also made significant ground before eventually fading into obscurity. Building a successful and sustaining social network application, it turned out, was not straightforward, as aptly demonstrated by Google's high-profile failure with G+ (launched in 2011). Getting a few million users was doable but gaining mass adoption on the global stage proved elusive for most of the decade.

The key technological breakthrough that allowed Social Media platforms to scale was the adoption of mobile technologies. These devices, particularly smartphones and tablets, put Social Media platforms into the hands, literally, of a global audience. By itself, the adoption of smartphones was significant, but what was crucial was that smartphones established a culture of always being connected. The iPhone was launched in 2007 and provided 24/7 access to social media applications so that one need never again miss out on a friend's update. This phenomenon allowed the next generation of social networking applications to leverage a human weakness that was previously unexploitable, known as the Fear of Missing Out (FOMO).

One of the contributing reasons for the phenomenal success of the new generation of social networks, such as Twitter, Facebook, and LinkedIn, was down to a complex combination of notifications and emails. These alerts encouraged users to check what was happening with their network constantly. Additionally, having learned their lesson from the MySpace debacle, they paid attention to the discretion and relevancy of advertisements shown on the platform.

FOMO is not something new; it has always characterized us. It is related to the fear that something will happen to people we care about and a primitive reflex that makes us feel good when we have everything under control. Social media companies expertly exploited this innate human dependency to create an addiction to the notifications and alerts on our phones. Users responded exactly as the platforms intended by constantly scrolling through their feeds to make sure they did not miss anything important.

Leveraging this dependency, the social media companies focused on keeping the users 'in-app' for longer. The new king of social media, Facebook, was quick to find a way to increase usage by including news in its services. In doing so, it marked the transition from social platforms to something different: a place to consume news. It worked very well. By July 2019, 52% of Americans received their news from Facebook, making it the most popular social platform for news sourcing (Suciu, 2019, para. 6). Some researchers have coined the term "Incidental news" to explain the dynamic whereby users get the news as a side effect of their constant connection to media platforms. "They encounter the news all the time, rather than looking for it" (Boczkowski et al., 2017, p. 1785). This is a shallow relationship with the news; partially read content, fragmentation, and exposure to only news sources or articles selected by invisible algorithms.

A users' Facebook became a seamless combination of editorial content right alongside a post written by a friend or acquaintance, creating a blurring of the boundaries between fact-checked and validated articles written by paid professionals held to account and content written by just about anyone. The algorithms would favor news and posts that were endorsed or commented on, which began to create digital echo chambers for users and polarization of views as their exposure to various views and arguments narrowed.

The danger caused by the polarization of content in Social Media is aptly explained by a well-known metaphor of German philosopher Immanuel Kant: if you were wearing a pair of rose-colored glasses everything in your world would appear rose-colored (Kant, 1929). The essence of the problem here is that users do not consciously choose to wear these "rose-colored" glasses. Some do not even know that they see content through a filtered process. Nevertheless, "rose-colored glasses" have been tailor-made for the user with great attention to detail and placed in front of their eyes.

This approach, which suits the business models of the current large Social Media platforms, creates great problems for society. It allows isolated or marginalized voices to appear as if they are mainstream, well supported, and even correct.

Prof Catherine O'Regan explains:

The Internet has also broadened the potential for harm. Being able to communicate with a mass audience has meant that the way we engage with politics, public affairs, and each other has also changed.

Hateful messages and incitements to violence are distributed and amplified on social media in ways that were not previously possible. (O'Regan, 2018, para. 3)

This amplification of marginalized voices spreading ignition, anger, and hate speech has been so prominent because it is precisely these types of posts that get shared, commented on, and liked the most on Social Media Platforms. What is beneficial for Social Media companies' profits is disastrous for societal harmony and individual wellbeing.

Author George Orwell in the book 1984, writes: "The horrible thing about the Two Minutes Hate was not that one was obliged to act as a part, but that it was impossible to avoid joining in" (as cited in Foley, 2016, para 3).

To paraphrase Orwell (1949), who controls the information controls the present, and who controls the present controls the future. To access and control the media means access to and potential control of public opinion (Jowett & O'Donnell, 2006). This idea was not lost on Facebook founder Mark Zuckerberg when, as reported by Foer (2017), he stated that "In a lot of ways Facebook is more like a government than a traditional company" (para. 11).

It is easy to see that Social Media was an untested, unchecked, and unregulated social experiment in hindsight. The harm of social media went unnoticed for many years. Like the proverbial frog in hot water, each new social media development was subtle and seemingly inconsequential. The public jumped on to the new platform with delight, and regulators could foresee no problems. Facebook, however, was not content with its level of success. It saw an opportunity to get larger, much larger. With this goal in mind, it hired psychologists and neuroscientists to learn how to make their application even more pervasive and addictive.

Their work paid off; people from across the planet now spend around 58 minutes per day on Facebook. In 2005, approximately 7% of American adults were active on social media. However, by 2017, that number had risen to 80% of American adults, and around the world, 3.5 billion people are active social media users (Kemp, 2020, para. 2).

There is a good reason why social media platforms want to maximize user engagement and time spent in-app. The more time a user spends on social media, the larger and more revealing their data footprint becomes. In 2015, researchers from the University of Cambridge and Stanford University released a study illustrating that after a user had clicked the like button more than 300 times, they knew them better than their spouse (Youyou et al., 2015). When user Max Schrems asked Facebook to give him all the data the company collected about him, he received a 1,200 page PDF. Facebook knows the power of data; it has fought with governments worldwide for the right to retain data forever on all users, even those that have deleted accounts.

MAIN FOCUS OF THE CHAPTER

Humans have always been a social species; in fact, communication and cooperation are human superpowers. However, since the turn of the century, we have been artificially amplifying how we interact with each other through technology. Machine Learning, a subset of Artificial Intelligence, has been deployed by large technology companies to make social media platforms more informative, stimulating, and entertaining for the users.

The problem with using Artificial Intelligence is that corporations deployed it to deliver higher levels of profit. The goal, therefore, became not a mission to widen the knowledge, awareness, and wisdom of humanity but to keep people on the platform. Furthermore, the most effective way to make social media 'sticky' is to make a user feel good. When a user receives a 'like' on their post, it stimulates the dopamine neurotransmitter, which triggers the brain's reward centre. The key to success was getting the algorithms to learn how to get user's posts liked just the right amount of times to trigger the chemical reaction. The 'like' button is probably one of the most successful and most adopted innovations in the history of our species. It gets pressed 4.5 billion times a day on Facebook, and every click provides more data to the platform and makes the app stickier. Social Media Platforms discovered that the posts that attract the highest levels of engagement demonstrate indignation or outrage. A 2018 MIT study found that on Twitter, from 2006 to 2017, false news stories were 70% more likely to be retweeted than true ones (Vosoughi et al., 2018). False news has greater novelty value than the truth and provokes stronger reactions — especially disgust and surprise.

ISSUES, CONTROVERSIES, PROBLEMS

It is possible that social media companies purposely engineered their platforms to predominantly feature posts that provoked strong emotional responses, often based on dubious content. For years they claimed they were not responsible for what users posted or how the information was shared, but their algorithms were likely tuned to promote fake news and shock stories because it manufactured likes.

The legal foundation for the modern Internet was Section 230 of the Communications Decency Act, known simply as Section 230: it outlined that websites and social media could not avoid being treated as traditional publishers, and as such, they would be considered legally responsible for the content they produce. US politician Nancy Pelosi and other members of the Democratic Party argued that tech companies should face harsher penalties for hosting misinformation or spreading hate speech (Feiner & Gram, 2020).

Social Media is fertile ground for sowing misinformation campaigns. During the 2016 US presidential election, Russia spread false information to more than 126 million people on Facebook and another 20 million people on Instagram (which Facebook owns) and was responsible for 10 million tweets. About 44% of adult Americans visited a false news source in the final weeks of the campaign. How was all this rendered possible? The answer is simple: advertising. Indeed, Russia bought Facebook ads that targeted and intended to inflame individuals based on race, gender, and partisanship (Bradshaw & Howard, 2018). It was necessary to know which profiles to show which fake news, rendered possible by former political consulting firm Cambridge Analytica. Cambridge Analytica specialized in "psychographic" profiling, meaning they used data collected online to create personality profiles for voters.

Cambridge Analytica was also responsible for helping Donald Trump to diffuse fake news during the 2016 US election. As reported by Illing (2018), the head of the Computational Propaganda Project at Oxford's Internet Institute found that "A disproportionate amount of pro-Trump messaging was spread via automated bots" (para. 9).

What is even more surprising and disturbing is that Cambridge Analytica was also able to analyze this information in real-time to determine which messages were most successful and in which geographic location and thus decide the next locations in Trump's propaganda agenda. This power to amplify and foment the opinion expressed through likes has given way to unprecedented targeted manipulation. In other words, if there were a spike in likes on a xenophobic article in a particular city, Trump would go there and give a speech on immigration.

The absence of regulation has made social media a wilderness of manipulation. Beyond the 2016 US election case, one of the most significant adverse effects of the spread of fake news is the erosion of trust and democracy. Social Media becomes a breeding ground for conspiracy theory, pseudoscience, and fake reports at the mercy of political propaganda or fake news.

The fact that dubious content spreads quickly and easily on these platforms implies how vulnerable people are to social media manipulation. The concept of cognitive bias was first introduced by researchers Amos Tversky and Daniel Kahneman in 1972 and it refers to an individual's systematic shift in judgment and understanding due to his/her personal beliefs and construction of reality (Kahneman & Tversky, 1972). Cognitive bias could manifest in various ways (there are more than 180 types according to the Cognitive Bias Codex). However, predominant in Social Media are false consensus, attention bias, and anchoring bias.

False consensus bias leads people to believe that their values and ideas are "normal" and that most people share these same opinions (Butler et al., 2015). Social bias creates a tendency to evaluate information from one's social circle as more reliable and as Morrison and Matthes (2011) demonstrate in their study, "individuals high in need to belong misperceive others' opinions on personally important issues as congruent with their own" (p. 712).

Social Media platforms leverage attention bias, which leads us to only pay attention to news and stories which confirm our opinions. Furthermore, the platforms limit our content to certain kinds of stories, ones we want to hear. These autonomous algorithmic decisions create a reinforcing loop, meaning that over time we will neglect more and more other topics, social battles, news, opinions leaving us in grave danger of believing that our interests and beliefs are universally shared.

This dynamic is called the filter bubble effect, and it is exploited mainly in microtargeting to successfully market products and services. It is a subliminal and potentially harmful type of advertising because it is invisible. For example, YouTube's algorithm autonomously decides what people watch on the platform 70% of the time (Solsman, 2018, para 2).

These are potent algorithms; 98% of Facebook's revenues are from companies advertising on their platform. Many companies provide their marketing spend to Facebook because it has more user preference data than any other company and exclusively shows products or services to users with a propensity to purchase. The number of repeat advertising customers for Facebook suggests that they are exceptionally proficient at transforming our data and metadata into profit. To quote the philosopher Sisto (2021):

From sex to food, from music to cinema, from reading to football, from politics to every form of social extravagance or peculiarity, there is no activity, situation, or passion that has not been recorded online (p. 47).

Simply put, Facebook records everything. There is even a folder in Facebook called 'offline activities' with information about users that third parties have decided to sell to Facebook. No totalitarian system has come to know the lives of its citizens in such detail.

Ironically, Social Media provides the illusion of a private space to share thoughts and ideas freely, but this is anything but true. Every action, every like, every share is recorded forever; even if a user deletes their account, the permanence of the data in other people's profiles remains. The Facebook policy states that "information that others have shared about you is not part of your account and will not be deleted when you delete your account". Facebook also states it will keep for a reasonable amount of time data recorded in its server. In Europe, the General Data Protection Regulation (GDPR) gives individuals the right to ask organizations to delete their data, but due to the hyper-connected nature of the internet, the right to be forgotten is more complicated than the simple demand of an individual for an entity to delete his/her data.

The evolution of social media data collection and behavioral-based advertisement targeting has been so fast that the average user can often still naively believe that the role of social media has remained that of socialization. Targeted advertising is a profitable strategy for companies, but that does not mean it is easy. Aguirre et al. (2015) have highlighted what is known as the personalization paradox, where broader personalization typically increases service relevance and customer adoption. However, at the same time, it can also increase customers' sense of vulnerability and thus result in lower adoption rates. As Bleier and Eisenbeiss (2015) demonstrated, personalized advertising could have the reverse effect by driving away consumers that perceive the approach as manipulative. For example, personalized advertising can spark serious privacy concerns by displaying intimate consumer information. Companies face an ongoing struggle to gain users' attention. Since personalized advertising needs relevant data, Social Media platforms must constantly innovate to find new means of collecting more personal data. Facebook may have exhausted the methods of extracting user data via their website and app since attention has turned to collecting personal data from new devices and ecosystems.

In 2014, Facebook announced that it had reached a definitive agreement to acquire Oculus VR, Inc., the leader in immersive virtual reality technology, for a total of approximately $2 billion. The acquisition, "positions Facebook to accelerate Oculus' growth in gaming, communications and new social experiences" (Facebook to Acquire Oculus, 2014, para 2.)

In 2020 Facebook made the headlines again when they changed the login settings on their flagship virtual reality headset Oculus Quest 2, to require all Quest users to log in with their Facebook profile. A blatant means of ensuring data collection was possible from the device. And what data they can collect! Commercial VR systems typically track body movements 90 times per second to display the scene appropriately, and high-end systems record 18 types of movements across the head and hands. Consequently, spending 20 minutes in a VR simulation provides just under 2 million unique recordings of body language for Facebook (Bailenson, 2018).

Facebook's Oculus studies gaze to determine preferences: a "heat map" of viewer data for 360-degree videos reports on what parts of a video people find most interesting. Oculus tracks biometrics and movement patterns with hand controllers; this data can record health and wellness data, such as whether a user is sick or tired on a specific day, along with health trends over time. Each headset also has externally facing cameras making environment analysis and facial recognition possible. Facebook is acutely aware of the vast amounts of new personal data available from a VR user and has prioritized the development of the Oculus platform. They recently announced the release of Horizon, a social experience where a user can explore, play and create with others in VR. This is a new world of possibilities and a world where Facebook owns all of the real estate. It is an advertiser and product placement dream. For these reasons, the Oculus hardware sells at near cost.

Google acquired Fitbit in January 2021, after two years of negotiations. A European Commission investigation concluded that google could not use Fitbit health data to inform a user's advertising profile; in other words, Fitbit data will be stored in a separate "data silo". Fitbit's privacy policy prohibits the sharing of any identifiable information to prevent things like insurance companies from denying patients health coverage or charging them a higher premium. However, these commitments have a shelf life of only ten years. Google can edit its privacy policy in 2031 to collect users' biometric data and share it with third parties.

As new technological devices are released, data collection proliferates. Most new cars are becoming 'smart, connected cars', equipped with dual high-definition cameras to record what is happening simultaneously inside and outside the car. Sensors and wireless technology transmit data to corporate clouds where artificial intelligence can learn about the vehicle, the driver, the passengers, and the surroundings.

Excessive data collection blurs the distinction between sensitive and non-sensitive data; it can open the door to discrimination, inequality, and restricted access to vital services.

More Issues, Controversies, Problems

The most successful social media companies have extraordinary abilities to extend the usage of their platforms and devices and extract user data without causing alarm. They have profited from what has been called surveillance capitalism. Successful social media companies do not charge users for access as they make their money downstream from providing targeted advertising.

Social media advertising accounts for 28.6% of all internet advertising revenues and will continue to grow. According to the 2019 Interactive Advertising Bureau report (PricewaterhouseCoopers, 2020, p. 5), revenues were $35.6 billion in 2019, a 23% increase over 2018 revenues of $28.9 billion.

Some social media companies do not use surveillance capitalism tactics; for example, a social media competitor to Facebook called Ello launched in 2014 as an ad-free social network. Ello is operating on more of a freemium business model, where a user pays for upgraded features that customize the experience. By 2019 however, Ello's revenues were just $5m.

An ad-free photo-sharing platform called Vero achieved similar revenues. What makes Vero unique is that it does not use an algorithm to rank its posts: all content appears chronologically. Another decisive feature is the policy of respect for users. As proof of this, Vero provides the app usage data to users to manage their screen time.

Additionally, its terms of use are explained in a user-friendly manner so that people understand what they are consenting to. The business model adopted by Vero is a subscription model: they offer a free subscription to the first 1 million users, and then they charge subsequent users with a fee. As honorable, however, as Ello and Vero's business models may be, they are making just 0.0058% of the revenues of Facebook. On the other hand, a new social media entrant called Clubhouse (launched in 2020) clearly states in its terms of service that it collects user data, stores it, and makes it available for Facebook to process. There is growing pressure in the EU to investigate Clubhouse's lack of adherence to the EU's General Data Protection Regulation (GDPR). Jaclyn Jaeger, Compliance Week, states (Jaeger, 2021, para. 1):

France's data privacy watchdog adds to a growing list of regulators that have launched investigations into Alpha Exploration, the publisher of the Clubhouse application, regarding measures it has taken (or not taken) to comply with the EU's General Data Protection Regulation (GDPR).

Despite rising tensions between consumers, regulators, and private companies about exploiting user data, Clubhouse has been a runaway success. In 2021, just 14 months after the launch, Clubhouse was valued at $4 billion. Scott Goodson, Author of Uprising exploring Movement Marketing says "If You're Not Paying For It, You Become The Product" (Goodson 2012, para 12). Clubhouse is free to use, and the

$4 billion valuation could indicate that the critical mass of users is not concerned about their personal data collection; instead, perhaps it is just a vocal few.

SOLUTIONS AND RECOMMENDATIONS

As consumer and regulator pressure builds on Social Media to modify data collection and processing conduct, some large companies will evaluate their future strategies. Google and Facebook, it seems, are committed to the centralized data collection model and will continue to build revenues in this manner. Facebook is betting big on data collection from Virtual Reality devices, and Google's launched Federated Learning of Cohorts (FLoC). According to Bindra (2021, para 4):

FLoC proposes a new way for businesses to reach people with relevant content and ads by clustering large groups of people with similar interests. This approach effectively hides individuals 'in the crowd' and uses on-device processing to keep a person's web history private on the browser.

FLoC is a way of personalizing content to users without collecting personal data or metadata; however, it can be considered just as invasive. For this reason, the Reddit, DuckDuckGo and WordPress websites are currently blocking FLoC.

The failure of Google to gain universal adoption to FLoC is a testament to how the public perception of centralized systems, data collection and data privacy is changing. In early 2021, Facebook experienced notable negative press from its WhatsApp privacy update. This update added tens of millions of users to competitors Telegram and Signal.

Brave a company launched in 2016 that provides a web browser, is experimenting with a new business model. The Brave website states (The browser reimagined, 2021, para. 1):

We all know what's wrong. As a user, access to your web activity and data is sold to the highest bidder. Internet giants grow rich, while publishers go out of business. And the entire system is rife with ad fraud.

Brave has based its business model on pay-to-view advertising and tipping for content creators using micro-payments. In return for viewing advertisements, users earn a small fee by being rewarded with a token called Basic Attention Token (BAT) that can be exchanged for products or provided as a tip for independent publishers or content creators. Users are not considered any more passive watchers for advertisements as they can actively choose where they want to direct their attention. Moreover, Brave locally matches ads to users without tracking or data collection, so personal data does not leave users' devices. The Basic Attention Token (BAT) is a cryptocurrency coin that utilizes the Ethereum blockchain. Cryptocurrency 'coins' such as Ethereum act as digital platforms for people to build a range of decentralized applications. These applications can include security programs, voting systems and methods of payment. Ethereum is more than an alternative to money; it can be used to codify, decentralise, secure, and trade just about anything.

Brave's business model is proving successful. In May 2021, Brave claimed 25 million users and 100% growth in the last year. Whilst this is just 1% of the 2.5 billion Chrome users, it can still be considered significant growth. The Brave browser intends to raise user awareness that the value of personal data and content is not shared equitably and provides what it considers to be a fairer system of distribution. In

2019 Twitter CEO Jack Dorsey announced that Twitter would build a new social media platform based on decentralized data called the Bluesky project. The new platform will include protocols and standards that allow users of different social media platforms to communicate. It will make it easier to enforce rules against hate speech and abuse. "The idealistic long-term vision is to make disparate social media networks more like email, so that users could join different networks but still communicate with each other no matter which one they're using," said Dorsey, as reported by (Palmer, 2019). However, critics of Bluesky point out that decentralized social network Mastodon has already implemented Twitter's ambitions to have less centralized control. Mastodon (launched in 2016) is known for being a friendlier kind of social network that keeps out the hateful or ugly content that proliferates on centralized networks. Journalists hailed it as "Twitter without Nazis'' (Jeong, 2017). Twitter, Facebook, and others have experienced issues with hate speech for several years, possibly because their autonomous algorithms have been promoting and amplifying shocking content, as described earlier in the chapter.

Whilst the Bluesky project has not made any significant announcements yet; others are moving quickly. Steem, a social blockchain that grows communities and makes immediate revenue streams possible for users by rewarding them for sharing content, has paid out over $59m to content creators. Steem is part of a new generation of platforms where the users become platform stakeholders, maintain control over their data, and earn cryptocurrency rewards for each contribution they make. The cooperative model may be as old as time, but the decentralized technologies making this possible on the web are very new.

The Steem community provides its members with a source of curated news and commentary and a means of getting high-quality answers to personalized questions. Steem has had operational challenges, and the value of its cryptocurrency token crashed; however, rising out of its ashes is a new social platform known as Voice powered by tokenization and created by a company called Block. One. Voice allows users to create digital arts that sell as unique digital artifacts. Their website says "The days of creators being exploited by platforms are coming to an end" (Zalatimo, 2021, para. 6). Block. One raised $4 billion from its Initial Coin Offering (ICO) and poured $150m into developing the Voice platform.

Equity and respect for personal data are at the heart of many new business models, and privacy-by-design is a core part of the strategy of several highly successful companies, including messaging app Telegram. CEO Pavel Durov believes that he can beat WhatsApp, a Facebook company, by winning on transparency and that maintaining consumer trust offers a significant business growth opportunity. Each time WhatsApp updates its terms and conditions, more and more users are ditching the platform to move to messaging applications that respect their privacy. Telegram currently has 500 million monthly active users, increasing 25% in less than a year. Other messaging apps like Signal believe that a Wikipedia-style donations system will help them to succeed in a climate where consumers are waking up to the significant downsides of giving away personal data to companies with only profit in mind. Signal and Telegram are unlikely ever to make the same level of revenues that Facebook does. However, the whole industry may have to accept smaller returns in the future if consumers are less willing to hand over their data and content for free.

FUTURE RESEARCH DIRECTIONS

New entrants to the social media market all believe they are solving 'the problem', and to each, the problem takes a different form. Clubhouse wanted to solve the Zoom fatigue and COVID-19 induced loneliness problems, so it launched an audio network. Mastodon wants to solve hate speech problems, Ello wants to solve the problem of surveillance capitalism, and Voice wants to solve the problem of creators not getting rewarded for their valuable content. Vero wants to reduce the polarization of views and digital content echo chambers, others want to introduce identity verification, and others want to reduce reliance on centralized control and decision making.

Two exciting developments that could turn out to be the technological trends that disrupt the Social Media industry are non-fungible tokens (NFTs) and decentralized autonomous organizations (DAOs). A non-fungible token is a unit of data that certifies a digital asset to be unique and therefore not interchangeable. NFTs can be used to represent pretty much any digital asset, from an image to a recording to a tweet. Decentralized autonomous organizations allow people to manage resources and make decisions based on minimal trust. DAOs have rules and bylaws encoded in a blockchain to ensure enforcement, and as such, all participants have assurance. According to Luis Iván Cuende, Cofounder at Aragon "Just as Bitcoin commoditizes storage and transfer of value, DAOs commoditize human coordination" (Cuende, 2020, para 12).

In the social media industry, developments typically occur in startups that raise money and try out a new idea. A new platform called social.network has been launched to enable user content to be created as non-fungible tokens (NFTs). Social. network's decentralized protocol does not extract value but instead allows individual participants to realize monetary rewards relative to the value they contribute to any given network.

NFT, (Non-Fungible Tokens) are becoming more and more popular. In simple words, they offer a way to represent anything unique as an Ethereum-based asset; for this reason, they have entered the scene through the Art Industry, gathering appreciation also in the gaming and collecting communities. Recently, they conquered the notorious platform e-buy, which allows the purchase and sale of these particular tokens.

However, NFT has shown to have even more significant potential. "Companies such as Unstoppable Domains and protocols like Namecoin allow users to purchase blockchain domains and even entire domain namespaces. The latter allows users to rent or sell individual domain NFTs down the road" (Chen, 2021, para 18). When a user claims a domain, it is embodied as an NFT on the Ethereum blockchain. The owner has full ownership and control of a decentralized domain. This network equity ownership is a real unparalleled disruption. Furthermore, the authors see compelling research directions in the possibilities unlocked by digital assets and, which give back to users their sovereignty. Indeed, "what ultimately disrupts many of the major web services created in the last decade could be peer-to-peer protocols, not companies" (Carlson-Wee, 2017, para 9).

One complaint about using cryptocurrencies as a reward model on Social Media is the environmental footprint of mining for the coins. Annually the mining for Bitcoin is estimated to consume as much energy as an entire European country. The CO2 impact of computer processing is both the advantage and disadvantage of Blockchain-based currencies. The huge compute effort to mine the coins is proof of work and creates significant environmental damage. To mitigate this, NET, the native token of the social.network, runs on a permissionless, Proof-of-Stake (PoS) consensus mechanism, which drastically reduces the energy requirements of securing transactions.

CONCLUSION

As Steem, Voice, social. network and other blockchain-powered tokenized social networks vie to be relevant, make it to the mainstream and achieve mass adoption; there are evident echoes of the past. Hi5, Orkut, MySpace, and Friendster were unable to scale due to the technological limitations of the era. Like these early networks, Steem, Voice, and social. network may feel like pioneers as they leverage decentralized technologies but may struggle to scale since these technologies are not yet mainstream.

Social networks that successfully protect their users' privacy, reduce hate speech, block bots, reward contributions, and respect personal data undoubtedly have a future. Many companies, however, have tried and failed to compete with the Facebook empire before. Many yet will try and fail. The success of a future competitor will most likely be down to their timing. In the next ten years, the Social Media industry will be a battlefield of corporations attempting to hold onto their fiefdoms, governments trying to limit tech companies' power, and consumers asking new questions about their data, their rightful share of the pie, and their rights. In 2030 Facebook might be the largest company ever built; it might host entire virtual worlds to spend your day, and your dollars, or perhaps social platforms may serve entirely different needs. However, since technologies and innovation follow consumer needs, the most prominent companies have to adapt. No company is ever too big to fail. Facebook is not too big to fail. Moreover, when they do, we can hope that we will not get a new king of social media; instead, the crown is discarded because there is no use for it in a decentralized, distributed, and equitable future.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Aguirre, E., Mahr, D., Grewal, D., De Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34–49. doi:10.1016/j.jretai.2014.09.005

Bailenson, J. (2018). Protecting nonverbal data tracked in virtual reality. *JAMA Pediatrics*, *172*(10), 905–906. doi:10.1001/jamapediatrics.2018.1909 PMID:30083770

Bindra, C. (2021, January 25). *Building a privacy-first future for web advertising*. Google Ads&Commerce Blog. https://blog.google/Products/Ads-Commerce/2021-01-Privacy-Sandbox/

Bleier, A., & Eisenbeiss, M. (2015). The importance of trust for personalized online advertising. *Journal of Retailing*, *91*(3), 390–409. doi:10.1016/j.jretai.2015.04.001

Boczkowski, P., Mitchelstein, E., & Matassi, M. (2017). Incidental news: How young people consume news on social media. In T. X. Bui, & R. Sprague Jr. (Eds.), *Proceedings of the 50th Hawaii International Conference On System Sciences* (pp. 1785-1792). Computer Society Press. 10.24251/HICSS.2017.217

Bradshaw, S., & Howard, P. N. (2018). *Challenging truth and trust: A global inventory of organized social media manipulation*. University of Oxford. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf

Brave Launches Next-Generation Browser that Puts Users in Charge of Their Internet Experience with Unmatched Privacy and Rewards. (2019). https://www.prnewswire.com/news-releases/brave-launches-next-generation-browser-that-puts-users-in-charge-of-their-internet-experience-with-unmatched-privacy-and-rewards-300957360.html

Butler, J. V., Giuliano, P., & Guiso, L. (2015). Trust, values, and false consensus. *International Economic Review*, *56*(3), 889–915. doi:10.1111/iere.12125

Carlson-Wee, O. (2017, January 8). *The future is a decentralized internet*. TechCrunch. https://techcrunch.com/2017/01/08/The-Future-Is-A-Decentralized-Internet/

Chen, D. (2021, February 24). *What are blockchain domain NFTs? A full introduction*. https://unstoppabledomains.com/blog/what-are-blockchain-domain-nfts-a-full-introduction

Cuende, L. (2020, June 11). *DAOs, the next big thing after social media.* Aragon. https://aragon.org/blog/daos-the-next-big-thing

Dickson, B. (2017, October 8). *Can blockchain decentralize the internet*. VentureBeat. https://venturebeat.com/2017/10/08/Can-Blockchain-Decentralize-The-Internet/

Facebook to acquire Oculus. (2014). https://about.fb.com/news/2014/03/facebook-to-acquire-oculus/

Feiner, L., & Graham, M. (2020, June 16). *Pelosi says advertisers should use their "tremendous leverage" to force social media companies to stop spreading false and dangerous information*. CNBC. https://www.cnbc.com/2020/06/16/pelosi-says-advertisers-should-push-platforms-to-combat-disinformation.html

Foer, F. (2017, September 19). Facebook's war on free will. *The Guardian*. https://www.theguardian.com/technology/2017/sep/19/facebooks-war-on-free-will

Foley, D. (2016, March 9). *Here's the original 'Two minutes hate'*. Intellectual Takeout. https://www.intellectualtakeout.org/Blog/Heres-Original-Two-Minutes-Hate/

Garrahan, M. (2009, December 4). The rise and fall of Myspace. *Financial Times*. https://www.ft.com/content/fd9ffd9c-dee5-11de-adff-00144feab49a

Goodson, S. (2012, March 5). If you're not paying for it, you become the product. *Forbes*. https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/?sh=3c309c6e5d6e

Illing, S. (2018, April 4). Cambridge Analytica, the shady data firm that might be a key Trump-Russia link, explained. *Vox*. https://www.vox.com/policy-and-politics/2017/10/16/15657512/cambridge-analytica-facebook-alexander-nix-christopher-wylie

Jaeger, J. (2021, March 19). Popular Clubhouse app being probed for GDPR violations. *Compliance Week*. https://www.complianceweek.com/gdpr/popular-clubhouse-app-being-probed-for-gdpr-violations/30181.article

Jeong, S. (2017, April 4). Mastodon is like Twitter without nazis, so why are we not using it? *Vice*. https://www.vice.com/en/article/783akg/mastodon-is-like-twitter-without-nazis-so-why-are-we-not-using-it

Jowett, G., & O'Donnell, V. (2006). *Propaganda and persuasion* (4th ed.). Sage.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454. doi:10.1016/0010-0285(72)90016-3

Kant, I. (1929). *Critique of pure reason* (7th ed.).

Kemp, S. (2020, January 30). *Digital 2020: 3.8 billion people use social media*. We Are Social. https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media#

Morrison, K. R., & Matthes, J. (2011). Socially motivated projection: Need to belong increases perceived opinion consensus on important issues. *European Journal of Social Psychology*, *41*(6), 707–719. doi:10.1002/ejsp.797

O'Regan, C. (2018). *Hate speech regulation on social media: an intractable contemporary challenge*. Research Outreach. https://researchoutreach.org/wp-content/uploads/2020/02/Catherine-O-Regan.pdf

Orwell, G. (1949). *1984*. Secker and Warburg.

Palmer, A. (2019, December 11). *Twitter CEO Jack Dorsey has an idealistic vision for the future of social media and is funding a small team to chase it*. CNBC. https://www.cnbc.com/2019/12/11/twitter-ceo-jack-dorsey-announces-bluesky-social-media-standards-push.html

PricewaterhouseCoopers. (2020). *Internet advertising revenue report*. https://www.iab.com/wp-content/uploads/2020/05/FY19-IAB-Internet-Ad-Revenue-Report_Final.pdf

Sisto, D. (2021). *Remember me: Memory and forgetting in the digital age* (A. Killgarriff, Trans.). Wiley.

Solsman, J. E. (2018, January 10). *YouTube's AI is the puppet master over most of what you watch*. CNet. https://www.cnet.com/news/youtube-ces-2018-neal-mohan/

Suciu, P. (2019, October 11). More Americans are getting their news from social media. *Forbes*. https://www.forbes.com/sites/petersuciu/2019/10/11/more-americans-are-getting-their-news-from-social-media/?sh=578f1ca43e17

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. doi:10.1126cience.aap9559 PMID:29590045

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(4), 1036–1040. doi:10.1073/pnas.1418680112 PMID:25583507

Zalatimo, S. (2021, April 13). *Coming summer 2021*. Voice. https://about.voice.com/blog/launching-summer-2021-nft-socialmedia-creators/

ADDITIONAL READING

Bolter, J. D., & Grusin, R. (2000). *Remediation: Understanding new media*. MIT Press.

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

Burrington, I. (2017, December 20). Could Facebook be tried for human-rights abuses? The Atlantic. https://www.theatlantic.com/technology/archive/2017/12/could-facebook-be-tried-for-war-crimes/548639/

Howard, G., & Katie, D. (2015). *The App generation: How today's youth navigate identity intimacy and imagination in a digital world*. Yale University Press.

Jones, K., Libert, K., & Tynski, K. (2016, May 23). The emotional combinations that make stories go viral. Harvard Business Review. https://hbr.org/2016/05/research-the-link-between-feeling-in-control-and-viral-content

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. doi:10.1126cience.aao2998 PMID:29590025

Mele, N. (2014). *The end of big: How the digital revolution makes David the new Goliath*. Picador.

Schneier, B. (2016). *Data and Goliath: The hidden battles to collect your data and control your world*. Norton & Company.

Sherry, T. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

KEY TERMS AND DEFINITIONS

Artificial Intelligence: The combined power of available data, algorithms based on learning approaches and computational resources.

BAT: Stands for Basic Attention; it is a token issued by Brave Company. On Brave Browser it is used to reward users for their attention, and allow them to tip content creators.

Bias: The inclination of the mind or a preconceived opinion about something or someone.

Blockchain: An immutable list of records, called blocks linked each other by means of cryptography.

Decentralization: The process of spreading control/authorities/data from centralised governance toward a distributed one.

GDPR: General Data Protection Regulation is a European regulation that governs how companies and other organizations process personal data.

ICO: Stands for initial coin offering, it is a means of crowdfunding that can help start-ups that want to meet short-term financial goals by issuing digital assets called cryptocurrency tokens.

Machine Learning: Is a branch of artificial intelligence (AI) and computer science which focuses on the use of algorithms to make inferences from patterns in data.

Retargeting: Is a form of online advertising which consists of showing users advertisements based on their previous browsing history on the firm's website.

Social Network: Websites that allow users to quickly create and share content with each other.

Surveillance Capitalism: A term coined by Shoshana Zuboff in 2014. It describes a market-driven process in which the goods for sale are your personal data captured through mass surveillance of the Internet by means of social platforms and search engines.

Token: A unit of value supported by blockchains. They only physically exist in the form of registry entries in the blockchain.

Tokenization: The process of issuing a blockchain token that digitally represents fractional ownership of a real tradable asset.

XR: Stands for extended realities, it refers to real and virtual world combined realities.

Chapter 11
Digital Ethics in Technology and Investments

Ritesh Jain
Independent Researcher, UK

ABSTRACT

This chapter focuses on aspects of digital ethics in technology and investments and covers the background of the growing surface area of the technology and its concerns related to data; it provides an overview of the data-related challenges and their ethical uses by organizations and people. It also covers emerging technology like artificial intelligence and its impact on ethical challenges. It provides an overview of the potential ethical challenges of technology that investors should consider and focuses on the criticality of the framework requirement and its implementation within businesses to make the right decisions. The author lays out the view on ethics and regulations and why companies should commit to ethical practices for their growth.

INTRODUCTION

Digital Technology is transforming our way of living, and this fast-paced transformation is causing ethical challenges for organizations and regulatory bodies globally. Growth in technology benefiting societies and people with a growing focus on Tech for Good. However, it's raising a fundamental question about being ethically digital.

Early growth in telecommunications and the internet led to connectivity & customer reach with low acquisition cost and incremental margins, which led to exponential growth in data per capita. Data and privacy regulations are still insufficient & vary significantly across the world. Big tech companies exploit and commoditize through social channels, eCommerce or logistics, and disruptive tech growth. Tech Giants like Google, Apple, Amazon, and Facebook hold significant amounts of customer data. They might not need to share the data per the regional regulations, and it's not ethical.

Technology involvement and the pace of its development across the businesses' value chains create a significant ethical dilemma for organizations trying to monetize through technologies and impact users. And it's a regulatory challenge for governing bodies. These challenges are well beyond organizations;

DOI: 10.4018/978-1-7998-8467-5.ch011

they create an ethical challenge for investors, as they need to analyze potential ethical issues, understand business models and risks. Regulators are struggling to respond to the pace of disruption and innovation set by tech giants. In the absence of defined regulations, it becomes the investors' responsibility to understand the potential issues and seek clarity. It is incumbent upon companies themselves to self-regulate. Global reach comes with global responsibilities and increasing risks to business models. Global platforms must maneuver local laws, geographical sensitivities, data privacy, and sovereignty with the lenses of digital ethics. It won't be an exaggeration to say disruptive innovators' primary risk relates less to business models than regulatory risk. Digital ethics in investment is becoming critical with growth in technology, investments, start-ups ecosystems, and disruptions. With start-up ecosystems and technology investments, lines blur between Impact investment, ethical investment, and digital ethics, which are perceived as a mirror of business ethics.

The Digital Ethics of Technology

Within a decade, and with a lot of hype, this transformation resulted in many platforms/ products and services being offered, leading to shifts in the traditional markets that were understood to be non-competitive, highly regulated, and regulated. The impact of this was not just a disruption to existing markets but created entirely new markets. The adoption and user-generated content in many ways democratized media and entertainment and had powerful implications for further regulation. Net neutrality is one such example; Net neutrality is the principle that all data on the internet should be treated equally, and open internet is essential for an open society.

Despite exponential growth in data generation, has the right balance of business practices and ethics been maintained? Only a few years ago, Facebook and Google were being scrutinized for data tracking and privacy concerns. Digital advertising has become the largest revenue generator for large and small companies, and it is a powerful marketing tool. However, who is the target audience for these services, and what data is collected? Should companies be able to collect data on the activity of users without their consent? With digital advertising, the company buying the ads need not disclose the primary purposes for collecting the data. One of the earliest controversies was about tracking web browsers without the user's permission.

Smartphone now has nearly 3.8 billion users (Turner, 2021), Facebook has two billion-plus users, over 1 billion WhatsApp users, and three billion-plus active Alexa users globally ("Global social media stats," 2021). Mobile payments are now a $4 trillion industry, and cash has not escaped the marketplace, with numerous mobile banking options and Alipay being by far the most dominant. On this backdrop, Tech for good (T4G) is beginning to debate by corporate organizations, policymakers, and their equivalents. T4G has been framed from the industry perspective as the ethical evolution of the marketplace. The transformation of the last century's technologies towards social good brought about the right conditions to create a virtuous cycle with breakthroughs in health care, information technology, and even media.

The Ethics of Data

Data was considered a neutral, informational value. However, this changed when algorithms became widely used on social media platforms and other online platforms, quickly generating large amounts of data. At the same time, questionable decisions may be made on data generated in the long term. Implications of a data breach are a considerable risk for companies and consumers alike. It impacts users financially to the extent they cannot access their personal and financial data. Data breaches are difficult to determine and remedy. It can lead to reputation and safety issues and consequential costs, including fines. Conversations around the value of data versus the confidentiality of data must occur for the industry to thrive (Jain, 2020b).

Data accumulation is a continuous process in business and can decide the future course of action for any business today. As such, business data strategies have become an essential ingredient for the success of contemporary corporations. Among many reasons is that data is a major driving factor in the digital transformation of any business. The payments space fared no different, and innovations have highly influenced it in data storage and interpretation.

Payment data is used by businesses like retail, financial services, e-commerce & manufacturers to make informed decisions and provide insight on forecasts for the future. Data can provide a competitive edge to a company to create & improve revenue generation streams and keep up with the competition. Moreover, it also provides value to consumers. Payment data is valuable in itself; however, it does not suffice in capturing entire consumer behavior. In the payments world, Data is categorized mainly in two parts – Personal Identifiable Information and Non – Personal Identifiable Information.

PII (Personal Identifiable Information) is one, which helps to identify an individual. Examples include: unique national Identification number, passport number, driving license number, bank accounts, email address, date of birth, and phone number, etc.

Personal data is typically further divided into sensitive PII and non-sensitive PII, or PII and non-PII. Sensitive data includes confidential information to an individual or a business, such as PII for health reports, contracts, or employer's recorded data. Non-sensitive Data or non-PII includes any available information in public records—for example, phone numbers in directories, business addresses, or tax account & VAT account numbers. The best way to differentiate between the two is that sensitive data would result in personal damages if lost or compromised. And so, it should always be encrypted.

Not to mention that when sensitive and non-sensitive data on an individual are critical ingredients of fraud and identity theft; for instance, one's bank account number, address, and phone number can be used for phishing frauds.

Sensitive, personally identifiable information such as personal information (Birth dates, Place of birth, Addresses, Religion, Ethnicity, Sexual orientation); employee number, tax ID, national ID; passport information; medical records; credit and debit card numbers; banking accounts; passwords; biometric information; school identification numbers and records; private personal phone numbers, especially mobile numbers, and email addresses. Non-sensitive information includes: business phone numbers and public personal phone numbers, employment-related information, IP addresses, cookies stored on a web browser. However, the categorization of data sensitivity varies across geographies.

Relevance of Spending Data

Another type of data that carries significant value is consumer spending data. Payment providers typically hold them. It provides powerful insight into consumer spending patterns, consumer behavior, preferred suppliers, regions, and demographics. Payment players use large chunks of data to gather insights on consumer spending behaviors. Moreover, correlating this Data with macro trends can enhance consumer offerings: from fraud detection to spending insights.

Data Sources

Data sources in business can be segregated across two segments. There will likely be a single data storage source in mature organizations, like a data lake, which holds Enterprise data. It is operational data like consumer preferences, needs, assessments, etc. The second type of business data source is anything external. Data collected externally is called supplemental data, and it can range widely. Some examples are data captured from external sources like digital ID's, social media, surveys, predictive modeling, sentiment analysis (Contextual Mining), etc. As mentioned earlier, payment providers get insight by combining supplemental data with their existing enterprise data. Several FinTech's and start-ups are also combining internal and external data for monetization.

Payments Data Enabling and Creating New Business

Payment players and third-party service providers have joined hands to monetize the payments data, for example, in consumer services. Cardlytics, a fintech company, has partnered with banks to provide consumer insight and geo-targeted marketing on mobile and online banking applications. Cardlytics rewards consumers for using their credit cards at specific retailers or locations. In a brief period, Cardlytics has partnered up with over 400 banks. And it now has more than 10 million consumers.

Financial fitness is another popular monetization channel. Take Plum, for example. It helps consumers make savings choices by capturing transaction data from their credit, debit, and current accounts. Plum then works to understand their behavior, using AI (artificial intelligence) to analyze their spending pattern and automatically transfer a certain amount into their savings account. In addition to B2C businesses, payments data have also enabled new approaches in B2B. Namely, SME Credit Lending and Risk management.

A consolidated view of SME data has enabled FinTech & Banks to build a more credit score on individual SMEs. They use non-traditional data such as the number of sales & purchases along with any outstanding receivable. This, in turn, supports credit lending for SMEs. SME Lending Platforms & Solutions like Provenir & Credolab empower fintechs and other financial services for credit decision-making, risk modeling, and lending.

In risk management, Kabbage is a great example. It implemented advanced analytics to get a risk score of online merchants. These scores are then used to provide working capital/credit lines to the small e-commerce businesses selling through online marketplaces. Kabbage also determines advanced approval and amounts by using data provided by the merchants in real-time. The data - such as merchant's revenue, net income, consumer traffic, and purchasing history combined with social network data - is also traceable online (Jain, 2020a).

Payment providers' data visibility - an advantage over others. Payment providers being in the middle of all the transactions of Banks, Telecom, Retail. It always has an upper edge when it comes to data monetization. Payment providers have less information on consumers but significant data on merchants in comparison to banks. However, their advantage truly comes to light when we consider that data monetization is most viable with two data - cardholders and merchants. Together, they give the full view of consumer transactions. There are opportunities like combining consumers with preferred merchants, channels, and potential products. Transactions to identify customer location, understanding the dynamics of markets at zip code level. Later we investigate how organizations are monetizing customer spending data and behavioral patterns.

AI in Payments and Retail

Artificial Intelligence (AI) has existed in financial services for some time and it's a fact that the technology offers significant benefits. AI increases efficiency reduces costs by optimizing processes and enhances customer experience. There are various use cases of AI applications from customer services, marketing, treasuries, portfolio management, trading, to name a few. AI is already augmenting or replacing human decisions in fraud prevention/detection/management, risk management, credit rating, wealth advisory, portfolio management. Financial institutions not deploying AI capabilities in these areas could be considered laggards. The amount of Data is ever-increasing and needs processing, and AI must improve operational efficiency and accuracy. With ever-growing & improving technology capability, the amount of available data grows, creating competitive pressures, the use of AI in finance is imminent. However, like any other new technology and its adoption, AI also brings various challenges; there are several challenges encountered by AI and concerning regulators, customers, and industry experts. These are categorized at a high level as: bias, accountability, and transparency.

The AI model is considered biased when its decisions could be against a specific segment of the population. It could sound like a rare occurrence to people with the simple assumption that machines are less judgmental than humans, but unfortunately, it is common. There are many incidents and stories from even the largest companies in the world.

In October 2019, researchers in the US identified a system algorithm used on more than 200 million people in hospitals across the US to predict the likelihood of extra medical care required for people, favored white over black patients. The race wasn't a variable in the algorithm. However, healthcare cost history was a variable that was correlated to race. The factor was that cost concludes how much healthcare requirement a particular individual has. Black patients incurred lower healthcare costs than white patients with the same conditions on average for various reasons (Paul, 2019).Thankfully, researchers worked with Optum to reduce the level of bias by 80%. But had they not been interrogated in the first place; AI bias would have continued to discriminate severely. Why do these biases happen? The main reason algorithms go, rogue, is incorrect problem framing; for, e.gAI systems checking the customer's creditworthiness are framed to optimize profits. It could get into predatory behavior and identify people with low credit scores to sell higher profitability products. This seems unethical, but AI wouldn't understand these nuances. Lack of social awareness is one of the reasons for unintended bias, e.g., if the data that exists in the system is biased and prejudiced that manifests the social system, machines neither differentiate nor understand these biases or can consider removing them. It just tries to optimize the model for biases in the system. Data representation is the key; if there is not enough data from a specific minority community and some of these data points are not great, algorithms could make generalized decisions based on the limited data. However, this is not much different from any human decisions influenced by available analytics.

AI is making its way through the payments and financial services world. Let's look into how AI is shaping the payments industry.

Firstly, AI in Debit Cards. Take Choice Financial as an example. They have partnered with Douugh to launch a debit card with an AI virtual assistant called Sophie. It will run diagnostics to enable Douugh to perform various tasks like bill payments and transactions while allowing the users to manage their savings and track their spending. (Ritesh Jain, Chris Mcloughlin, Fred Bar, Kim Ford, 2020). Another widespread usage of AI is in Banking Chatbots, which is becoming common. Bank of America was the first major bank to implement an AI chatbot, Erica. The bot helps customers manage their finances through voice and text interactions to resolve queries. It also advises on financial services. However, the area in which AI is looking to create the most impact is Fraud Detection. With the growing volume of digital payments, the need for real-time fraud detection is imminent. AI brings scale and speed to fight against payments fraud. Payment frauds are growing and becoming more sophisticated and complex. They vary in digital footprint, patterns, and structure. These variations mean that they could go undetected by rules-based logic and predictive models alone.

AI-Based fraud detection addresses both complexity and speed. For instance, the fintech firm Kount provides an AI-powered fraud detection system with the fastest response rate for calculating risk scores. At 250 milliseconds. Its fraud scores are also twice as predictive as other methods (Kount, 2019). On the opposite side of fraud, AI can also help to reduce "false declines". MasterCard, for example, is planning to facilitate actual real-time authorization to prevent false declines. A common disturbance in the customer's shopping experience. Other big techs such as Amazon and Google are also working on something similar. But how else can we prevent false declines? Comes in the innovation in Payments Authentication. These are the Biometrics technology, such as Voice and Image recognition.Biometrics technology alone is already a precious tool for payment processing and authentication. It addresses the multi-factor authentication challenges. Biometrics requires a scan of fingerprints, faces, eyes, and voices. Apple Pay and Zwipe (in cards) are some of the use cases today (Jain, 2021). Biometrics like fingerprints, iris, and faces are unique and safer than passwords, pins, and cards, which could be compromised. However, from the customer experience perspective, the main advantage of biometric authentication is that consumers always have the prerequisites required to complete transactions. A world with just a biometric scanner would be a great place. One wouldn't have to carry any additional physical devices. Finally, not only does biometric authentication reduce false declines, but fraud is almost impossible with biometrics authentication. Especially with multiple factors involved, there is no technology available to replicate retina AND fingerprints yet. Let's now take a further look into conversational commerce. This is when businesses can be interacted with through their messaging and/or chat apps. But how are conversational commerce methods fueling payments? Well, AI and Natural Language Processing (NLP) are making science fiction a reality. Payment cards are critical to online commerce, and by extension, conversational commerce. The majority of users use their Credit and debit card accounts with online retailers and other online purchases. The most preferred method of using these cards for online transactions is to enter the card information at checkout; the second most preferred method is using a card on file with the e-commerce merchant. Conversational commerce is leading to Intelligent Authentication, such as Voice-Based Authentication. This, in turn, leads to innovation and opportunities for fintechs and New businesses. Some examples may already be familiar to us, and people can already shop with their voice via Alexa, Siri, and google home.

Data Monetization

Let's now discuss some use cases of data monetization. WeChat, a Chinese super app, turns the payments industry practices on its ear by handling payment processing within its ecosystem. It has generated billions of dollars in annual revenue from transaction processing. Additionally, WeChat also utilizes payments data to gain insights. They then use these insights to improve customer experience and identify additional revenue opportunities. There are multiple similar Superapps in Asia and in Africa. Phonepe and Paytm In India, Gojek in Indonesia, and Opay in Africa, to name a few. We have said time and time again that businesses are using payments data to generate insights and revenue. But let's see now what exactly payment service providers and networks are doing with their data.

Consider Mastercard, one of the most recognizable payments service providers. Mastercard Retail Locations Insight (MRI) is an analytics service that generates insight based on billions of transactions occurring globally daily within the Mastercard network. These solutions analyze key variables like location, incomes, social & demographic aspects for accurate decision-making globally. Businesses can use this information to evaluate upcoming area development by those in charge of a specific area's urban planning and economic development. This solution also helps the companies to make decisions based on critical variables related to a location. Decisions like where to open a new marketplace or buy and sell commercial properties - thus leading to new opportunities.

Mastercard's competitor, Visa, also has something very similar. Visa's analytics platform provides data-driven insights, which are powered by VISA's global payment network. It provides easy access to three years of Visa debit, credit, and prepaid data across the Visa portfolio. The data can help businesses analyze performance at the cardholder, merchant, and transaction levels. More importantly, this does not mandate any technology integration.

Payments data is also vital for government, health, and other public sector organizations to plan for supply chain, change in regulatory and compliances, forecast, and planning, during normal times, and it becomes of utmost importance during times of pandemics and national emergencies. World Bank, World Health Organizations, and World Economic Forums utilize this Data for their planning. Traditional players like VISA and Mastercard can directly monetize their data by setting up Data as a Service platforms.

Predictive Analysis in Banking and Retail

Finally, before we conclude this topic, we cannot mention AI in finance without predictive analytics. Taking center stage in Banking and Payments, Predictive Analytics has seen significant growth in the last couple of years and is expected to grow at twenty percent year on year. And the main reason for this growth is acceleration and sophistication in money laundering, payments & card fraud.Customers interact with banks and financial services through multiple channels leading to significant growth in data. This data is utilized to gain customer insights and predict their behavior leveraging AI & ML technologies. Predictive analytics help financial institutions enhance efficiency and reduce cost by optimizing their processes & products. Let's discuss the application of predictive analysis. By reducing false-positive and detecting fraud, predictive analytics identify the minute difference in the transactions to segregate legitimate from fraud transactions. They are reducing card default rates by identifying cardholder behavior and predicting the likelihood of paying their debts. Predicting customer behavior to forecast and model customer lifetime value and relationship with banks or financial institutions.

Impediments in Data Monetization

So far, we have discussed the new payments business model and how data can be monetized in payments. Nevertheless, there are multiple challenges in data monetization which are (Jain et al., 2020): business ethics and privacy, technology, regulatory guidelines and compliance.

Business Data Ethics and Security

First, let's discuss business data ethics and privacy. The Cambridge Analytica scandal has spotlighted data privacy and how some companies collect and monetize consumers' data. Striking a balance between consumer protection and data monetization is a challenging task. Companies must consider both the quality of data and consumer privacy, businesses need to be aware and adopt these three factors about monetizing consumer data while remaining compliant with regulations and safeguarding consumer data privacy.

Value of Business Data to Others

To maximize the value of their data, data-collecting businesses should recognize that anonymous and aggregated data are key contributors in making investment decisions. Personally, identifiable information has zero value beyond being a legal liability. For example, the number of soft drinks a company has sold across Europe has more relevance than how many soft drinks an individual named" John" the consumer has consumed. Moreover, enrichment and aggregation also enhance the value of data. For instance, a global logistics company working with different vendors across the country to deliver soft drinks can track soft drink sales and provide insight to their vendors as a value-added service.This would be targeted to help vendors promote and improve their sales according to their locations or insights based on a particular product sales figure.The logistics company can monetize this data gathered over a while.

Technology Challenges

The second set of challenges that we will discuss are technology challenges.

Some of the typical roadblocks in obtaining helpful insights are (Jain et al., 2020): First: too much data everywhere. Widely distributed data is a common impediment to data monetization initiatives. Second: data access. This is a common roadblock from a consistency and formatting perspective; getting data in a consistent and useful format is the challenge. Third: data cleansing. This is an infinite challenge that will exist as long as an enterprise creates new data. The challenge lies in the fact that multiple data sources tend to get contaminated along the way. Lastly: data scalability. This is critical not only from the collection and storage (warehousing) perspective but also from consumption and processing speed, access, and security. The issue grows with volume, variety, and velocity.

It is of utmost importance to know that data is handled across organizations and how they are held accountable for whether it is at rest or in motion. We can divide the data handling process into three stages: Data Disclosure, Data Manipulation, and Data Consumption. Within each step, the company must be ethical at all times, but how? Data is collected in real-time or sourced from archives or backup during the data disclosure phase, and data should only be retained, transformed, or disclosed given the relevant consents. It must also be encrypted throughout. Then, in data manipulation, data is aggregated and transformed using multiple data sets or APIs from various providers. The resulting Data, having been manipulated, must remain secure and audit-worthy. Finally, in data consumption. Data analytics can be context-driven and need intention-based consent for data usage from users and contextual consent from businesses extending associated APIs.

Regulatory Challenges or Opportunities?

The Global Data Monetization market is growing over twenty percent year on year, and it's due to an increase in the volume of data and reduction in data storage cost. However, the varying structure of global regulatory policies is a challenge in data monetization. New regulation policies are being implemented, like the General Data Protection Regulation (GDPR) in the European Union (EU). It empowers EU citizens with data ownership, and organizations need to amend their approach to customer data and monetization (European Commission, 2018). Businesses like British Airways and Marriott have been charged with hefty fines for data breaches, and regulators are getting tougher on companies to ensure compliance. Facebook was fined five billion US dollars in the Cambridge Analytica case to breach customer trust and data privacy. Similarly, WhatsApp and Twitter also faced data privacy fines. Nevertheless, there are also opportunities. Data is being generated on a large scale from various data sources, Fitbit, Google, or Facebook.Businesses resort to data monetization technologies and tools to create revenue streams and create insights for decision-making.Data is growing and available from various sources. It is a challenge as most unstructured Data is often unutilized; data monetization tools help integrate data from multiple sources and refine unstructured and structured data. Yet, the usage of any data must accompany stringent compliance practices, among other challenges.The Quality of Data is of utmost importance as low-quality data could lead to unexpected outcomes. The quality of data within the organizations impacts the choice of the data monetization tools, limiting the growth of the data monetization vendors.

Ethics in Investing

Historically, ethics are not the first to be considered when it comes down to investments directly related to money-making. The primary function of businesses is to make money for shareholders, and even for the investor's primary reason to invest is to get better returns. However, this landscape is changing with the awareness, and a new generation of more ethically minded investors, businesses, FinTech's, are capitalizing on this. The leadership of the organizations or the government is accountable to ensure tech companies' societal impact, purpose; however, Investors are another critical group who plays a pivotal role.

During the last decade, specifically in recent years, responsible investing has become mainstream. Whether institutional investors, venture capital, private equity, or even angel investors, they consider integrated environment, social, governance (ESG) factors into their selection and management of their investments. ESG practices are picking up pace and ingrained across the asset classes. However, venture capitalists are still struggling to have a systematic approach for screening, managing ESG performance, and another societal impact of the technology investments. Various standards and government bodies are supporting these through multiple initiatives. Environment, Social and Governance (ESG) is a driving force in the new setup, and businesses are measured on these factors for sustainability and societal impact. Historically ESG factors lagged behind financial and political for business and financial decisions; however, its changing and ESG are critical factors in making these decisions. Enterprises realize the benefit of keeping ESG is a significant factor. Market data shows that investing in businesses and assets with a higher ESG rating leads to better returns with lower risk. Morningstar reported over a decade, the average annual return for a sustainable fund invested in large global companies has been 6.9% a year compared to 6.3% with a traditionally invested fund (Bioy, 2020).

There are various approaches for ethical investing, and they have a lot of common factors. Most commonly, by investors/or investment managers avoid the companies that seem problematic, companies involved in high-risk businesses like firearms, tobacco, alcohol, fossil fuels. And another approach is to focus on companies with a direct, measurable impact on the environment and society, e.g., businesses that focus on green/ alternative energy, businesses/ retailers who only engage with suppliers following fair trade policies. Other approaches are faith-based, like Catholics, Muslims, or Vegans. Some of the investors are focused on action and support by the businesses for various causes like women empowerment, empowering minorities, human rights, climate change.

Despite the growth in ESG/ Social Impact investing, there is a significant gap in people looking for these investments and advisors and options available for such investments. As we've discussed, technological advancement is beneficial. However, this growth and opportunities create ethical dilemmas for companies in monetizing using these technologies, impacting the users, and government bodies are overwhelmed with regulating the business activities. Hence, the responsibility comes down to investors to analyze the ethical issues when engaging with companies and understand their model efficiency and risks. One needs to consider the variety of ethical impacts on users and the environment, like discrimination, consumers' choice, privacy, societal impact, waste. And to assess these, a robust framework required focuses on corporate culture, business functions, governance, disclosures and practices, regions of operations, and market/social data.

Technology and Ethics- Investor's Considerations

Investors like venture capital, private equity play a significant role in the growth of the future of technology, in turn, society. Investors are the first to support and provide a platform for even many of the world's largest companies or big techs, including Google, Facebook, Amazon, Twitter, Airbnb, Uber, and many other leading tech companies globally. Venture capital investors play a critical role in influencing business model, organization values, and culture being the initial investors and foundation of values. Culture is shaped in the initial years of development, venture capitalists being the investors and board members play a vital role in the process. Many leading tech companies have faced significant challenges in managing societal issues and scrutinizing their foundational value and culture by governments, regulators, and media. Like Google and Facebook faced increased scrutiny in recent years for their business model, revenue generation for selling user data, crushing their competitors, and impacting human rights. And other companies faced challenges regarding ensuring safe working conditions of employees, living wages. As we've discussed, technology advancement is beneficial; however, this growth and opportunities create ethical dilemmas for companies in monetizing using these technologies, impacting the users and government bodies overwhelmed with regulating business activities. Hence, the responsibility comes down to investors to analyze the ethical issues when engaging with companies and understand their model efficiency and risks associated with it. One needs to consider the variety of ethical impacts on users and the environment, like discrimination, consumers' choice, privacy, societal impact, waste.

Discrimination

Big techs and social media companies have prohibited targeted advertisement for demographics. However, they could still advertise jobs to specific age groups, which is considered discrimination. Concerns related to Artificial intelligence are still there and can exacerbate, create unintended biases. Discrimination is a significant challenge and addressing this will require more than technology. Machines can't be held responsible for making decisions as decisions are based on the data and its representation; there are many incidents, whether it's related to the creditworthiness of men vs women or ethnic people or insurance, medical attention for ethnic groups. Technology can't be neutral, and you can't abstract away social context.

Customer Choice

The network effect can drive market concentration and reduce competition which impacts customer choice. Big Techs have intensified this concern through their business strategies using dominance in a sector to capture other markets and sectors. With deep pockets, they can easily manage the cost, regulations, and compliances, which provides them with an unintended advantage over their smaller and younger competitors and leads to market concentration. Competition market authorities (CMA) globally focus on such concentration. Open Banking in the UK is one of the prominent examples of it. CMA identified market concentration in the banking sector and improvised it through PSD2 to address innovation and create a competitive market.

Privacy

Privacy is a significant concern with growth in data and customer channels and should be an essential factor for investors while considering technology investments. Data privacy is responsible for collating, using, and storing customers' data according to their expectations, meeting regulations and compliance, and following data ethics. Amazon echo logs every interaction, mobile devices track user's physical movement, and social media like Facebook enables us to track friends. All these tools are helpful. However, they empower big techs to target users with granularity for their growth; some regions are less data privacy-focused than others, like China and Russia, which could lead to human rights violations. On the other hand, the European area is stringent with data privacy policies and introduced GDPR (General Data Protection Rights) to empower consumers.

Social Impact

The technology investment landscape is changing due to environmental, social & governance (ESG) factors; sustainable development goals (SDG) initiatives. Society now expects businesses to be responsible; companies need to focus well beyond their operating model sustainability; they need to consider the social impact, value creation for long-lasting effect on society. Work-life balance for employees for their overall wellbeing, businesses carbon footprints, waste generated by a business, business's impact on customers overall wellbeing. Investors also seriously consider social impact while making investment decisions. Technology-focused on positive social impact reap benefits from customers, investors, regulators, and government authorities. ESG factors are becoming essential for businesses as this represents a socially responsible business and attracts investments and better returns. And to assess these, a robust framework is required, focusing on corporate culture, business functions, governance, disclosures and practices, regions of operations, and market/social data.

Corporate Culture

An ethical environment and culture where employees are empowered to voice their concerns about culture, products, ethics of the organization instead of focus on productivity, empowered to do the right thing. The organization's culture represents people's ethical behavior, working environment, and management. Businesses need to build an environment where employees can focus and balance disruption and innovation without ethical challenges. Companies need to have the process to monitor day-to-day operations and any potential ethical issues.

Organizational Governance

The organization's governance structure is critical to maintain and enforce ethical behavior. This includes processes, policies, board structure, board experience, and expertise in identifying ethical issues that other people might be oblivious to, involvement, and oversight of ethical problems. Independent directors and their involvement, how often they discuss ethical challenges in the business model. Directors' background in identifying and assessing potential ethical issues within the organization or technology—a record of discussions of ethical issues, potential impact, and remediations at the board level.

Disclosures and Practices

Organizations with robust disclosures and practices could enhance their competitive advantage, reduce potential ethical issues, and gain confidence from customers and investors. There are multiple examples of large businesses how businesses handled or mishandled disclosures. e.g., Microsoft has broad disclosure practices and strict policies for their own business and their customers to use Microsoft technology in targeting customers. On the other hand, Facebook faced the press for the Cambridge Analytica scandal. IBM decided to stop selling general-purpose facial recognition technology with their overall commitment to their values and supporting opposition of all forms of human rights discrimination. Disclosures and practices showcase organizations" ability and commitment in balancing their ethical dilemma in monetizing information with the responsibility to protect their consumers. Disclosures are part of regulatory and nowadays organizations voluntarily disclosing and publishing their information and procedures, which various indexing and benchmarking organizational use.

Regions of Operations

As mentioned above, certain regions and countries are likely to capitalize on user data or take strict actions against businesses that do not cooperate, creating risk and business opportunities. Generally, a region with low levels of human rights can create contrary, however positive outcomes for the local businesses, which could be uncomfortable for large global businesses which require higher ethics and privacy standards to operate in those regions. By benchmarking and index-providing a ranking of state of privacy in the regions, countries are a reliable source to investigate before making investment decisions.

Applying an Ethical Framework

Ethical behavior can be guided by laws, regulations, standards, and codes of ethics; however, individual judgment is critical for making ethical choices and conducting appropriately. Individuals and businesses must have a well-developed framework of principles to refer to during ethical dilemmas; otherwise, their thought process can lead to indecision or fraudulent conduct, leading to image and public trust destruction. Establishing an ethical framework is a crucial step for businesses to engage in ethical behavior. Venture capitalists, private equity, and investors are used to making decisions from a business's profit and loss perspective; given the priority of ethical behaviors in their professional responsibilities, they must analyze any potential ethical issues or challenges. Implementing a framework and guidelines for making decisions helps investment professionals analyze their conduct from conflict of interest to professional obligations.

The Relationship Between Ethics and Regulatory Guidelines

Ethical behaviors are often seen as equivalent to legal behavior. Regulations and laws put appropriate constraints on businesses' natural tendency to pursue, which could harm the interest of others. Regulations guide businesses and people towards ethical behavior; however, regulations do not cover all aspects of unethical behavior. Ethical and legal behaviors are distinguished as what is required by law and morally correct. Ethical behaviors, principles go well beyond the legal boundaries; it's about what is the right thing to do even when no one is watching. Specifically, with all the innovation and disruption, regulators can't keep with the pace. They lack sufficient resources to enforce well-conceived regulations; hence relying on rules to establish ethical behavior is challenging; thus, businesses must ensure ethical behavior of technology and investments to create an ethical culture in the industry. There is much evidence that individuals or businesses succeeded at circumventing the regulations for personal gains; robust ethical framework implementation could limit the abuse at the personal or business level. On the other hand, regulators need to ensure that they demonstrably and consistently adopt the highest ethical standards.

Commitment to Ethics by Businesses

A business's code of ethics risk is primarily ignored if it is not inherent in the business processes. The ethical decision-making framework enables businesses to bring aspirations and principles of the code of ethics to life beyond the compliance exercise, which is the core of the business culture. People's natural inclination to do the right thing must be augmented by building a culture of ethics and integrity.

Senior management in the organization is responsible for developing, maintaining, and leading by example for a strong culture of integrity and ethics.

Adopting a code of conduct that imbibes the ethical principles that define the organization's intent and expectations from its employees is a critical initial step. And let's be realistic. Ethical behavior and practices can't be just guided through the policies and code of conduct. Organizations need to continuously educate ethical decision-making and reinforce, monitor, and empower employees to question potential unethical issues.

CONCLUSION

The digital revolution is causing a change in the way individuals, communities, organizations, governments, and even private and public sector organizations conduct their business and function. Digital technology creates and impacts how everything we do is transacted, purchased, marketed, and researched. The emergence of new possibilities raised new challenges regarding its ethics and policies and how organizations and regulators address those challenges. The evolution of digital technology is expanding the ethical and governance-related issues that require multi-stakeholder partnership in addressing. It is also driving the need for individuals, communities, and organizations to think and act ethically. Investors are responsible and accountable to look for Digital and Business ethics while making investments in new businesses, whether existing or start-ups.

REFERENCES

Bioy, H. (2020, June 16). *Do sustainable funds beat their rivals?* Morningstar. https://www.morningstar.co.uk/uk/news/203214/do-sustainable-funds-beat-their-rivals.aspx

European Commission. (2018). *Data protection in the EU*. https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

Global social media stats. (2021). *Datareportal*. https://datareportal.com/social-media-users

Jain. (2020a, December 15). *Fintech Future.* SPD Salford Professional Development - Gov Tech UK.

Jain, R. (2020b, Dec 21). *Payments strategies and business model innovation (1.0)* [Online course]. Centre for Financial Technology and Entrepreneurship. https://courses.cfte.education/payment-strategies-business-model-innovations-course

Jain, R., Mcloughlin, C., Bar, F., & Ford, K. (2020, Dec 21). *Payments in digital finance 1.0* [Online course]. Centre for Financial Technology and Entrepreneurship. https://courses.cfte.education/payments-in-digital-finance-specialisation

Kount. (2019, June 26). *Kount launches next-generation AI, changing how payments fraud prevention is delivered*. https://kount.com/announcements/kount-launches-next-generation-ai/

Paul, K. (2019, October 25). The healthcare algorithm used across America has dramatic racial biases. *The Guardian*. https://www.theguardian.com/society/2019/oct/25/healthcare-algorithm-racial-biases-optum

Turner, A. (2021, August). *How many smartphones are in the world?* BankMyCell. https://www.bankmycell.com/blog/how-many-phones-are-in-the-world

KEY TERMS AND DEFINITIONS

Big Tech Companies: Leading technology companies like Google, Apple, Amazon, Facebook, Microsoft.

Data per Capita: Per capita is a term used in economic and statistical analysis that means per person; per capita information provides more granular data than just aggregate information; Data per capita means data generated per person.

Data Privacy and Sovereignty: Data privacy is the right of a citizen to control how personal information is collected and used. The concept of data sovereignty is closely linked with data security, cloud computing, and technological sovereignty.

Tech for Good: Tech FOR Good is a community of people, projects, organizations, and funders promoting the role of technology to improve social, environmental, and economic outcomes.

Chapter 12
Business Ethics in a Digital World:
A 360 Perspective

Ingrid Vasiliu-Feltes
University of Miami, USA

ABSTRACT

This chapter will highlight the importance of transforming our conceptualization of business ethics in the digital era and the opportunities related to an optimal design of sustainable digital business ethics programs in this new hyper-connected, hyper-automated digital world. The complex issues of this revised business ethics model will be addressed from three perspectives: corporate governance, leadership, and society. The sections related to corporate governance will highlight the operational challenges when aiming to incorporate ethics into the boardroom's DNA and will emphasize the sustainability imperative ethical business leaders are facing in this digital era. This chapter will also posit that by adopting a design thinking approach for business ethics in this digital era, we can leverage all the benefits offered by emerging technologies and scientific advances while maintaining a human-centric stance.

INTRODUCTION

The digital era has produced novel technologies, created a whole new socio-economic ecosystem, triggered new legal and regulatory requirements, gave birth to a new global workforce, redefined competitive advantage, shifted the risks and rewards balance, and has offered us new opportunities.

We are currently experiencing the 4th industrial revolution and are likely entering the 5th. Although this transition will offer exciting opportunities and will likely have a profound global transformative impact it will likely exacerbate current digital ethics challenges and create new ones we never envisioned before.

Business Ethics in the digital era poses some of the more complex and nuanced challenges and several publications have highlighted the wide range of ethical questions that remain to be answered in addition to a whole new set of ethical challenges that will arise while we leap into the next industrial revolution.

DOI: 10.4018/978-1-7998-8467-5.ch012

Blending boundaries between physical, digital and biological worlds will likely continue at an exponential pace with technologies such as AI, IoT, 6G or next generation computing having the potential to reshape, recalibrate or disrupt our society and the global economy.

Leaders that wish to be prepared for this leap into the 5th industrial revolution and successful in managing a new digitalized hyperconnected global workforce will be required to display a complex armamentarium of novel skills, such as technological fluency, embracing of design thinking methodology and mastery of applied digital ethics. It used to be sufficient to be a “tech savvy” leader, however with the current exponential advancements of a variety of modern technologies it is imperative for business leaders to have not only a higher degree of technical acumen in order but also emotional intelligence, lead with purpose and embrace ethics in order to remain competitive, as well as to ensure long term sustainability for their companies. Additionally, with the increase in consumer-centricity and a broader acceptance of *technology-as-a-service* business models current leaders would be advised to accept and manage change and be willing to disrupt themselves in order to avoid being disrupted by others. Digital data governance, proactive ethics- and compliance programs, ESG-consciousness and a focus on United Nations Sustainable Development Goals are becoming an expectation for the new generation of leaders. Versatility in foundational ethical concepts and willingness to adopt complex ethics frameworks at every step of technological deployments are an imperative in this new business environment defined by a high degree of digitalization, virtualization and by intelligent automation of business processes.

There is an ongoing debate about the responsibilities of the business community and the incentives or mandates required to enforce these. We must balance moral theories with the practical corporate social responsibility. It is incumbent upon us a society to rethink how those moral theories can be applied mindfully and efficiently in this digital era.

Complex digital ethics framework include an ethical and digital risk analysis, a set of tools that can be deployed at various stages and in different layers of the business processes, establishing and monitoring digital ethics implementations via Key Performance Indicators, as well as building a culture of digital ethics while embracing an ongoing quality improvement mindset. Business leaders usually struggle with deployment of these frameworks due to a lack of digital ethics literacy and due to a failure to harmonize with other key organizational units such as Compliance, Human Resources, Marketing to only mention a few.

There continue to be many myths related to ethics which are only exacerbated when aiming to address digital ethics. Furthermore, we continue to encounter a profound misunderstanding about the impact of digital ethics on all aspects of an organizational ecosystem. Last but not least, we have observed ongoing resistance from some business leaders to embrace a proactive approach and they seem to continue to prefer a reactive crisis mitigating approach to ethics breaches.

A best in class cyber-ethics program likely represents one of the major drivers of success when operationalizing proactive digital ethics frameworks. More often than not cybersecurity and ethics are managed in silos and therefore offer the opportunity for cyberattacks or privacy breaches. A state of the art comprehensive digital ethics program must always include a robust cyber-ethics component and be fully integrated or harmonized with the organizational cyber-security systems.

Ethical marketing is one of the domains that has often been fully neglected and has recently gained increased attention. Digital marketing has been plagued by several key ethical challenges such as mindful data usage, targeted advertising, truthful advertising and refraining from adverse comments related to religion politics, ethnicity or other emotionally charged content.

Lastly, it is crucial for business leaders to understand that compliance and ethics are different disciplines despite several aspects that require marked alignment in policies and processes.

BACKGROUND

Ethical values have been in existence since ancient Greece and their application to individuals and society has sparked countless debates for centuries. However, the application of ethical virtues to the business ecosystem and a more formal approach to business ethics is relatively novel by comparison. Although there are numerous definitions of business ethics all of them clearly outline that they can be applied at an enterprise, individual and societal level. Business ethics in the digital era only adds further layers of complexity and therefore has been intentionally avoided by many corporate leaders. There are still very few organizations that have a state of the art ethics governance structure, a proactive ethics program and that include creating a culture of ethics in their strategic roadmap.

The scientific and business communities, numerous not for profit- and government agencies are all appropriately concerned about ethical issues unique to this newly emerging hyperconnected hypervirtualized world, with digital ethics hot topics like bias, discrimination, privacy, data transparency and trust making the headlines almost on a daily basis and have therefore caused increased societal awareness.

Corporate Ethics in the Digital Era

Before we outline some of the complex elements related to deployment of corporate ethics in the digital era, we must ask ourselves a more basic question: do our current business practices meet even the most basic ethics principles? Do corporate business practices uphold justice and are they meeting the impartiality or equality conditions? Beneficence and Non-maleficence; although these two can certainly be addressed with appropriate governance the latest technological advances have identified our cyber-vulnerabilities and the need to further enhance our cyber-ethics efforts. Confidentiality, fidelity, integrity; artificial intelligence algorithms, neural networks, adversarial networks, decentralized autonomous organizations are producing large datasets products or solutions that are vulnerable to cybersecurity attacks and we will have to design robust proactive cyber-defense mechanisms to protect our privacy, enhance trust and maintain data integrity. Autonomy; this principle usually sparks controversial and passionate debates with no conclusive answer.As this cursory overview of the most basic ethics principles highlights, the answer is not as simple in this digital era and self-sovereignty remains aspirational at this point.

There is an opportunity to create shared value within our society by capitalizing on social challenges and transform them into new business opportunities. The key differentiating factors between a state-of-the-art strategic business plan and a state-of-the-art corporate ethics program are essential to understating the potential long term value for our society (Menghwar & Daood, 2021) more complex approach describes the various nuances that differentiate digital ethics, digital business ethics and digital ethics leadership. Digital ethics addresses behaviors related to digital mediums, norms related to the use of digital tools, autonomy an ownership of online data, etc. Business ethics includes governance, social and fiduciary responsibilities, as well as discrimination, fraud, abuse or bribery. Leadership ethics describes the attributes of ethical leaders in this digital world.

Ethical Leadership in the Digital Era

Ethics conscious business leaders are expected to display a high regard for moral values such as honesty, fairness, respect for others. It will be important for business leaders to dedicate resources to all of those and to embed them into the fabric of their enterprise strategy. By striving to demonstrate ethical leadership in this digital era, business leaders can greatly contribute to the development a Global Business Ethics Culture. As Heraclitus stated: “Character is Destiny “.

Leading by example and cultivating external awareness are crucial particularly during times of crisis or disruption. Additionally, an ethical leader shows empathy, puts his team first and displays a high degree of moral integrity.Digital data ownership, digital identity, digital privacy are the foundational elements and blatant violations keep making headlines on a frequent basis. At a more advanced level the fusion of AI, IoT, DLTs, next generation computing, next generation networks (6G and beyond) will create the need for increasingly complex digital ethics and cyber-defense programs. It has become evident that exponential adoption of emerging technologies and ongoing innovative forces will require a sustainable digital ethics culture in order to prevent potential violations of trust, integrity, digital privacy, digital ownership. Business leaders should be appropriately concerned about general ethical issues, as well as those unique to each industry. Reactive or mitigation approaches are not an optimal solution and one would hope that ethical leadership in the digital era will be marked by state of the art strategic planning and a careful tactical deployments. A great quote that illustrates the importance of creating a robust ethical governance infrastructure within the organization was by Alfred Einstein who stated “ relativity applies to physics, not to ethics”.

Impact of Digital Ethics on Society

To underscore the importance of social impact, the authors of a recent publication on business ethics (Islam & Greenwood, 2021) noted the need for a double movement in their journal—one toward focusing specifically on ethics and a broader inter- and cross-disciplinary approach. They highlight the crucial aspects that are typically addressed in normative ethics, while also calling attention to the more complicated elements that cause frequent expert debates such as solidarity, justice and empathy.

MAIN FOCUS OF THE CHAPTER

Digital Ethics has been defined in numerous ways by experts. Most definitions aim to capture the complex aspects related to moral principles or rules of engagement that govern our digital interactions. Individual aspects, societal dilemmas, informed consent, data governance and data monetization are all important domains. A recent article also acutely highlights the need for heightened transparency in qualitative research and introduces us to the ethics-as-a-process approach (Whiting, R., & Pritchard, K., 2017). Digital Ethics has an impact on the complete 360 spectrum of our global business ecosystem and in this chapter we wish to highlight the importance of proactive digital data governance, the need for a new generation ethical leadership suitable to the demands of this digital era and the inevitable transformative impact on our society.In order to successfully deploy any proactive business ethics programs organization need to complete a through ethical risk analysis, embrace emerging technologies, integrate their digital transformation processes with the rest of their enterprise strategy and embed ethical tools and checklists

Carefully balancing individual moral values versus corporate social responsibility is essential in creating a successful sustainable ethics culture. In the digital era leaders that master the art of business ethics understand this complex intricate relationship and can harmonize it with their digital strategic plans. So what are some of the key characteristics defining ethical leaders in the digital era and what toolkits are they employing most often?

Digital fluency, agile management style, higher risk tolerance, versatility in sustainability strategies, embracing diversity and inclusion are a few of the traits embodied by ethical leaders in this digital world. Having a robust corporate ethics governance structure, building a culture of digital ethics, serving as an example of upholding the individual code of conduct and deploying a proactive digital ethics program are all part of the ethical leadership portfolio.

We would like to distinguish between the ethics of individual behavior which focuses on the individual's actions and the ethics of conditions which focuses on the circumstances under which the individual is supposed to act. Many of our current ethical concepts were developed on the foundational elements derived from premodern societies, which reflected a totally different socio-economic landscape and were in stark contrast to modern conditions. The ethical benefits of markets and competition are complex and warrant significant attention (Lutge & Uhl, 2021).

While experts define business ethics as the application of ethics to business, they also highlight that a more complex level it can be understood as the study of good and evil, right and wrong and just and unjust actions of business leaders. "Business ethics is the study of appropriate business policies and practices regarding potentially controversial subjects including corporate governance, insider trading, bribery, discrimination, corporate social responsibility, and fiduciary responsibilities. The law often guides business ethics, but at other times business ethics provide a basic guideline that businesses can choose to follow to gain public approval."(Alexandra Twin).

In a recent paper published we are introduced to a framework for how to manage business strategy and ethics. The author provides a complex overview of how a state of the art enterprise strategy can be aligned with a proactive ethics program, as well as how to incorporate social resopinbility into the corporate action plan (Moran, 2021).

Issues, Controversies, Problems

Regulatory, compliance and legal considerations have always ranked high on the list of challenges that have prevented a large scale deployment of emerging technologies. They are also frequently one of the main challenges encountered when designing and deploying large scale digital ethics programs. Numerous international organizations have collaborated and dedicated their efforts to developing frameworks to address the barriers inherent to cross-border business ethics globally.

Bridging the gap between ethics frameworks and successful deployment, as well as maintenance of robust ethics programs is essential. ESG-consciousness, diversity, and inclusion are all equally important drivers of success and ethical business leaders have to balance these against the financial pressures they are facing. The rapid pace of digital innovation and transformation also increases the cyber-security risks exponentially and digitally savvy business leaders will understand the need for cutting edge cyber-security and cyber-defense programs.

Emerging Technologies continue to challenge ethicists, business leaders, compliance and regulatory decision makers in adapting to the novel dilemmas they bring to light. Examples include: high frequency networks such as 6 G and 10 G, Internet of Things, satellite internet, artificial intelligence tools, distributed ledger technologies, decentralized autonomous organizations, digital assets, digital twins, 3 D bio printing, bio-implants, brain-computer interfaces, next generation sequencing techniques like CRISPR, next generation computing technologies (such as neuro-morphic, DNA or quantum). The challenges span from the ethical deployment of these technologies across industries to deciding if to use these technologies for their business operations deploying them to actually perform certain functions that can lead to increase efficiency and shareholder value. In this fast paced, high risk and high return digital era C suites must consider transparency, explainability, diversity, inclusiveness, alignment, bias, privacy violations, self-sovereignty etc. (Hochheiser & Valdez, 2020).

Diversity has been one of the core focus points for business ethics advocacy and the digital era has only caused an exponential increase in awareness. However, the opinions are split among experts regarding the impact of digitalization, automation, globalization. There are those who believe that the digital divide has worsened diversity efforts and those who believe that it offers unique opportunities to reduce the diversity gap. Furthermore, socio-economic and cultural factors also deeply impact the ability to design and deploy effective digital ethics programs and many organizations globally are lagging in their diversity and inclusion efforts. The recent pandemic has exacerbated the need for diversity at all levels of society and within organizations (Gilshan, 2021)

Regulators, scientists and industry leaders will have to collaborate to develop best practices for large scale technology deployments, and argue in favor of an ethics-by-design approach to prototyping (Bourgais & Ibnouhsein, 2021).

A futurist could imagine the plethora of ethical issues related to a "Digital CEO" with a baron-computer interface implanted, communicating with the C-suite via a cloud-based neuro-morphic computing platform, using satellite internet, who also has bionic eyes implanted and runs a corporation that sells autonomous vehicles, using 10 G and IoT networks. This corporation could also deploy smart contracts and embed advanced artificial intelligence tools to improve performance of their product, increase their costumer base and for their internal human resources functions.

Corporate social responsibility (CSR) policies are evidenced by adopting socially relevant business practices for people, communities, companies, and related institutions. They eloquently describe the complex relationship between establishing policies and practices to encourage standardization of social responsibility within organizations and the positive impact they found in their Chilean survey (Arenas-Torres, & al, 2021). In another example, gender differences in 38,179 individuals from 36 countries in 9 relatively homogeneous global regions were analyzed. The ethical fillets derived from gender egalitarian views and their impact on workplace harassment were evaluated in depth by the authors of this published study. Their findings confirmed that regardless of significant regional differences, there is a higher incidence for women to feel subject to harassment and a significant correlation with other indicators used for gender equality rankings. One of their conclusions is that women that have the opportunity to live in a country or region that offers better economic opportunities and where women leadership is promoted will perceive less workplace harassment. The study's findings also showcase how important moral and societal values are in driving gender equality. Furthermore, the authors found how a discrepancy between advocacy and actual implementation of gender equal best practices in organizations can actually lead to increased perception of harassment (Otterbach & al, 2021).

Another body of work published recently aimed to assess the impact of corporate communication strategies. This initiative collected and analyzed data from Facebook pages of the Top 100 Global Brands, the authors. Their findings were significant as they confirmed the importance of clear and transparent communication at a corporate level. The authors also evaluated several indicators, shared methods for benchmarking and highlighted the impact of external stakeholders on the impact of internal corporate social responsibility communications. (Yang & Basile, 2021).

A multi-vocal literature review published illustrates how ethical leadership work has changed in the last few decades from 1985 to 2020. The findings of this study that reviewed 83 bodies of work highlight the important impact technology advancements had on leadership characteristics and how they influence ethical and moral behaviors of leaders that had to adapt to the new societal and business demands (Bhatta, 1970).

Another important paper has investigated how the alignment of two corporate functions, sustainable supply chain management (SSCM) and trade compliance (TC) can help companies to take corporate value chain responsibility (VCR). The authors aim to tackle the complex topic of corporate value chain responsibility and highlight evolutionary theory nuances. They also point out potential dilemmas and challenges when we aim to design corporate strategies (Baier, 1970)

Organizational behavioral management is essential in this digital era as it has the potential to profoundly influence the success of proactive digital ethics programs. In a recent publication on applied behavior analysis the author emphasizes the most important asset we have when aiming for increased return on our investments: the people. Professionalism, team work, practice standards and upholding moral values must be addressed simultaneously to achieve long term success. (Weatherly, 2021).

SOLUTIONS AND RECOMMENDATIONS

1. Design Thinking

Business leaders of the digital era will have to harmonize digital innovation, digital transformation, digital ethics and design a new sustainable human-centric ecosystem. There are several methods available to accomplish this, however the application of design thinking principles seems uniquely suited for a highly dynamic and rapidly evolving global digital economy. The adoption of design thinking has rapidly transcended industries and is currently widely used by business leaders. There are seven core design thinking stages: empathy, definition, ideation, prototyping, selection, implementation, and feedback. It allows for increased speed of implementation, improved user satisfaction and cost savings. So how would a design thinking powered business ethics playbook encompass? The most important and likely most difficult to accomplish is a change in culture by shifting to a human-centered mindset, encourage creative confidence in all employees and the leadership team, express empathy for all stakeholders, and embrace uncertainty. It will be essential for business leaders to complete an honest digital ethics and business risk analysis, as well as to incorporate the voice of the customer in order to create a robust, resilient and sustainable culture of digital ethics. By encouraging all employee to brainstorm, ideate and submit their proposals during this process it will be easier to implement and maintain the data governance program at an enterprise level, as well as to uphold an ethical code of conduct for employees. By adopting an agile prototyping model that allows frequent iterations and embedding of ethical principles in the DNA of each of these minimal viable products. Only by receiving ongoing feedback from employees regarding the ethical deployment of technologies can businesses achieve increased adoption and a sustainable culture of ethics. Throughout this process it would be essential to maintain transparency and open communication with all employees as that will facilitate engagement, acceptance and successful deployment. After these stages would be accomplished the focus would shift to crafting a state of the art implementation plan in order to ensure long term sustainability and revenue generation. As stated, design thinking is only one option in the armamentarium available to business leaders. However due to the emphasis on user experience and outcomes, its ability to empower diverse teams and a process that encompasses an ongoing innovative loop it is this author's opinion that design thinking should be considered as a methodology of choice during this crisis and beyond. Through the convergence of inquiry, empathy and cultural change in digital technology deployments a design thinking methodology can address some of the major challenges in the current business ecosystem (Hamington, 2019). Business leaders that aim to be socially responsible would be well advised to embrace design thinking when embarking on their digital innovation or transformation journey. (Brown & Wyatt, 2010.)

2. Ethics as a Competitive Advantage

Ethical issues are not often mentioned in C-suite environments and almost never seen as a competitive advantage, however the digital era is likely to change this paradigm. Experts foresee that robust business ethics programs are an expectation and impact investing will be one of the predominant instruments for upcoming generations. Furthermore, social support of the organizational brand and a social purpose will likely transition from aspirational statements to become core business strategies. The corporate ethical values scale will certainly be profoundly impacted by the marketing practices that have emerged during the digital era, while brand credibility will play a key role (Zayyad, 2020).

When deploying digital business ethics programs there are a few key drivers of success Thorough gap and risk analysis, robust data governance programs, embedded ethical guardrails, monitoring ethics KPIs, develop and embrace a culture of digital ethics are a few essential ones. Business leaders must consider ethics deployment as important as all other enterprise strategic goals. Ethical risks have to be recognized, managed or mitigated. Leaders have to instill a culture of integrity through modeling in their own behavior. An open and transparent culture is encouraged at all levels within the organization and ESG-consciousness balances the high return expectations from shareholders.

FUTURE RESEARCH DIRECTIONS

1. Culture

There are major cultural differences that can significantly impact upholding of ethical principles from an individual, societal and business ethics perspective. Regardless if one has a cultural relativist or ethno-centric approach, digital ethics and business ethics are difficult to harmonize. Given that is expected to have an ongoing increase in the global multi-cultural hyperconnected workforce we are likely to witness an accentuation of the cultural and digital business ethics divide.

The dual relationship between culture and ethics has fascinated us for centuries. The nuances related to cultural norms have a major impact on the design of proactive effective and efficient business ethics programs. Intercultural ethics is a sub-field that warrants our attention and dedicated research given the hyper-connected and hyper-virtualized society we live in where traditional geographical boundaries have been blurred. Having a pluralistic ethical approach when designing technology implementation playbooks will be essential to a successful large-scale adoption and to leverage technology for the greater social benefit (Aggarwal, 2020). The digital era has only exacerbated these cultural aspects due to a global workforce and the applied digital ethics discipline has become increasingly complex requiring a broad spectrum of skill sets. Some experts are even calling for the need of a Digital Ethics Officer role. This signals an awareness about the complexity of the challenges we encounter, as well as the ongoing nature of creating a digital culture of ethics. Deploying ethics is a not a static, one time project and actually required a continuous quality improvement approach as new products, new services, new business partners, new software will always require a careful ethics assessment. For centuries ethical principles have been foundational to our society, however with the rapid adoption of emerging technologies we are currently witnessing an ethics revival within the scientific and business community due to the complex ethical issues we are facing in this digital era. Business leaders have the opportunity to shape the future by fostering a Culture of Digital Business Ethics and by contributing to the development of a Global Digital Business Ethics Ecosystem.

The latest pandemic has once again accentuated the signifiant ethical crisis we experience as a society and the high propensity to not uphold human rights during these challenging times. A recent study conducted aimed to analyze Covid-19 impact on business ethics in the Indian economy in the commercial world with business and remedies by implementing ethical standard in business during the pandemic situation (Surthi, 2020)

2. United Nations Sustainable Development Goals and Human Rights

One of the global agendas that has the potential to align a large group of key stakeholders is the attainment of United Nations Sustainable Development Goals. Regardless of the cultural or socio-economic differences most organizations and governments around the world can coalesce at least around one of the SDGs. Creating proactive digital business ethics programs has a positive exponential effect by facilitating achievement of the United Nations Sustainable Development Goals agenda. There are also strong opinions from experts that recommend including SDG protections in the business code of ethics and developing safeguards to ensure that we don't violate global privacy rights in our efforts to improve other SDGs. From a business ethics perspective SDG leaders will have to carefully balance the environmental and social dimensions with the economic pressures. Upholding justice, being inclusive and respecting diversity will be paramount in accomplishing ethical attainment of the SDG 2030 agenda (Salamat, 2016).

It is essential to understand the factors that impact of socially responsible governance on corporate revenues. A recent study found that there were marked differences during crisis versus non-crisis time periods, as well as a string correlation between other societal dimensions such as health of the economy, degree of trust or upholding human rights (Tsai & Wu, 2021).

Several experts have called attention to the fact that we have expanded the duties traditionally assigned to states and governments to businesses who are now expected to be equally accountable in upholding ethical virtues and human rights. However there continues to be disagreement regarding the extent of accountability and how to harmonize this additional role with their corporate duties as expected by shareholders (Brenkert, 2016). Furthermore, a recent publication suggests that international human rights law could play an important role in shaping business strategy and guide the development of a digital culture of ethics (Nersessian, 2018)

Equally important is the change in our approach to graduate education. If we do not embed the core sustainability principles into the academic curriculum for business schools we can not expect a different generation of business leaders that will be accountable and execute the United Nations agenda. A recent editorial shared a framework for how ethics, corporate social responsibility and management education can be harmonized, and accelerate attainment of the sustainable development goals (Setó-Pamies & Papaoikonomou, 2020). On a broader political economy scale, it would be helpful to recalibrate our teachings to incorporate foundational ethical values and build a robust digital ethics culture. Future generations of global citizens must have a higher degree of digital ethics literacy to mitigate against privacy breaches, mass surveillance and loss of data ownership (Mills, et al., 2017).

3. Nouveaux Digital Ethics

We are likely on the cusp of the 5th industrial revolution, which will likely bring about even greater ethics challenges due to the increased blending of digital, physical and biological boundaries as well as the exponential adoption. One way to mitigate potential negative consequences would be to introduce digital ethics into the education system, as well as in to the corporate training and development processes. The reactive model to only conduct an ethics risk assessment when there are identified breaches is not sustainable in the digital era. Many experts also suggest that certain specific elements could be required instead of recommended as they would further enhance the successful implementation of digital ethics, such as:diversity, inclusion or ESG mandates.

We must also take into consideration that new technologies will also trigger novel ethical dilemmas. Bionic humans, bio implants, digital placentas, digital twins, brain-computer interfaces, human-animal hybrids are only a few examples that cause significant ethical dilemmas even currently. Only two decades ago brain-computer interfaces would have been mentioned solely by futurists or science fiction authors (Jawad, 2021). Further research will be required to analyze the intricate relationship between culture, society norms, legal barriers and moral values when deploying cutting edge technologies such as brain-computer interfaces (Burwell & al, 2017).

The types of ethical questions that arise can span from the most basic such as questioning if certain emerging technologies should be even allowed or not, to the more nuanced and complex business ethics issues such as digital identity, neuro- or biohacking, genetic data privacy, allowing robots to teach or perform human resources roles, if we can have a CEO with a brain-computer interface, or if we can have international business transactions completely run by a DAO (S*ulkowski, April, 2020).*

Smart Cities development brings tremendous excitement and has brought together numerous international stakeholders to design frameworks, strategic roadmaps and playbooks. This "smart cities revolution" aims to rethink, recalibrate and rebuild urban developments to me more resilient, sustainable and affordable. However, not enough attention is being paid to embed ethics into the fabric of these smart cities (Chang, 2021). Furthermore, the massive data sets being generated via smart cities deployments also cause significant ethical debates among experts related to adequate stewardship of those datasets. A recent article questions if civic data governance is a feasible solution, highlights some of the risks associated wit that approach and concludes that even civic projects are not immune to abusive practice by those who aim to monetize data in unethical ways (Artyushina, 2020). The issue of data privacy is further explored in a book chapter about technological sovereignty for better data care in smart cities. The authors emphasize the tremendous potential for technology deployments to create more livable, sustainable and equitable cities, while also cautioning us about the dangers of ethical violations. They propose a new model that can serve as a blueprint and aims to improve data literacy, trust and autonomy. (Foth & al.,2021) In order to mitigate the potential negative consequences that could emerge from deployment of emerging technologies that would infuse all our society and be embedded in the DNA of our business processes future generations of ethicists will have to embrace change, tolerate disruption and engage in a culture of continuous improvement to adjust to the dynamic digital needs.

CONCLUSION

In order to adapt and thrive in this digital consumer-centric era businesses must understand, accept and embrace business ethics. Most experts agree that companies that earn the consumer's trust have higher returns, however that will also require changes in strategic choices, mindful decision making processes, valuing reputation and social impact. They must integrate ethics into every fiber of their strategic plan, collaborate with industry stakeholders to encourage ethics and social responsibility in their extended business ecosystem and invest in building a lasting culture of ethics. The soft ethics concept highlights the need for us to strive to go above and beyond regulatory and compliance requirements. As a society we must have foresight and dedicate ourselves to both hard and soft ethics in order to adapt to the demands of the digital era (Floridi, 2018). Almost a decade ago an article called attention to the fact that the digital era has the potential to violate human rights and suggested the development of a Declaration of Digital Rights (Mathiesen, 2012). Given the lessons learned from the last few years and the recent pandemic it is imperative to develop a Universal Code of Conduct for Digital Business Ethics and a Global Digital Ethics Framework. As digital ethics advocates we must hold ourselves accountable in order to build a digital ethics legacy for future generations.

REFERENCES

Aggarwal, N. (2020, September 19). Introduction to the special issue on Intercultural Digital Ethics. *Philosophy & Technology*. https://link.springer.com/article/10.1007/s13347-020-00428-1

Arenas-Torres, F., Bustamante-Ubilla, M., & Campos-Troncoso, R. (2021). *The incidence of social responsibility in the adoption of business practices*. MDPI. https://www.mdpi.com/2071-1050/13/5/2794

Artyushina, A. (2020). *Is civic data governance the key to Democratic Smart Cities? the role of the Urban Data Trust in sidewalk toronto*. Telematics and Informatics. https://www.sciencedirect.com/science/article/pii/S0736585320301155

Baier, C. (1970). *Challenges in the implementation of responsible business conduct*. OPUS 4 | Challenges in the implementation of responsible business conduct. https://opus4.kobv.de/opus4-ku-eichstaett/frontdoor/index/index/docId/674

Bhatta, N. (1970). Emerging ethical challenges of leadership in the digital era: A Multi-Vocal literature review. *Electronic Journal of Business Ethics and Organization Studies*. https://jyx.jyu.fi/handle/123456789/74932

Bourgais, A., & Ibnouhsein, I. (2021). *Ethics-by-design: The next frontier of industrialization*. AI and Ethics. https://link.springer.com/article/10.1007/s43681-021-00057-0

Brenkert, G. G. (2016). Business ethics and Human Rights: An Overview. *Business and human Rights Journal*. https://www.cambridge.org/core/journals/business-and-human-rights-journal/article/business-ethics-and-human-rights-an-overview/4E12322863D6BA2B17871B03EDA9BBB9

Brown, T., & Wyatt, J. (2010.). *Design thinking for social innovation*. Development Outreach. https://elibrary.worldbank.org/doi/abs/10.1596/1020-797x_12_1_29

Burwell, S., Sample, M., & Racine, E. (2017). *Ethical aspects of brain computer interfaces: A scoping review*. BMC Medical Ethics. https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-017-0220-y

Chang, V. (2021). *An ethical framework for big data and smart cities*. Technological Forecasting and Social Change. https://www.sciencedirect.com/science/article/abs/pii/S0040162520313858?dgcid=rss_sd_all

Floridi, L. (2018, February 17). *Soft ethics and the governance of the Digital.* Philosophy & Technology. https://link.springer.com/article/10.1007/s13347-018-0303-9

Foth, M., Anastasiu, I., Mann, M., & Mitchell, P. (2021). From Automation to Autonomy: Technological Sovereignty for Better Data Care in Smart Cities. In B. T. Wang & C. M. Wang (Eds.), *Automating Cities. Advances in 21st Century Human Settlements*. Springer.

Gilshan, D. (2021). *The ethics of diversity*. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2021/02/03/the-ethics-of-diversity/

Hamington, M. (2019). Integrating Care Ethics and Design Thinking. *Journal of Business Ethics*, *155*(1), 91–103. doi:10.100710551-017-3522-6

Hochheiser, H., & Valdez, R. S. (2020). *Human-Computer interaction, ethics, and Biomedical Informatics*. Yearbook of medical informatics. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7442500/#:~:text=Ethics%20in%20Human%2DComputer%20Interaction,and%20accountability%2C%20among%20others%201%20

Islam, G., & Greenwood, M. (2021). Reconnecting to the social in business ethics. *Journal of Business Ethics*. https://link.springer.com/article/10.1007/s10551-021-04775-7

Jawad, A. J. (2021). Bioethics of medical devices based on brain computer interfaces (BCI). *Journal of Clinical Research & Bioethics*. https://www.academia.edu/45585122/Bioethics_of_Medical_Devices_Based_on_Brain_Computer_Interfaces_BCI_

Lütge, C., & Uhl, M. (2021). *Business ethics: Interdisciplinary perspectives and future challenges*. Oxford Scholarship Online. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198864776.001.0001/oso-9780198864776-chapter-8

Mathiesen, K. (2012) *Human rights for the Digital age*. Taylor & Francis. https://www.tandfonline.com/doi/abs/10.1080/08900523.2014.863124

Menghwar, P. S., & Daood, A. (2021). Creating shared value: A systematic review, synthesis and integrative perspective. *International Journal of Management Reviews*, ijmr.12252. Advance online publication. doi:10.1111/ijmr.12252

Mills, K. A., Stornaiuolo, A., Smith, A., & Jessica Zacher, P. (Eds.). (2017). *Handbook of Writing, Literacies, and Education in Digital Cultures* (1st ed.). Routledge. doi:10.4324/9781315465258

Moran, W. (2021) *Technical career Institutes, Inc., (TCI College) strategic analysis*. CUNY Academic Works. https://academicworks.cuny.edu/qb_pubs/72/

Nersessian, D. (2018). *The law and Ethics of Big Data Analytics: A new role for international human rights in the search for global standards*. Business Horizons. https://www.sciencedirect.com/science/article/abs/pii/S0007681318301095

Otterbach, S., Sousa-Poza, A., & Zhang, X. (2021). *Gender differences in perceived workplace harassment and gender egalitarianism: A comparative cross-national analysis*. Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1111/beer.12338

Salamat, M. R. (2016). *Ethics of sustainable development: The moral imperative for the effective implementation of the 2030 agenda for sustainable development*. Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1111/1477-8947.12096

Setó-Pamies, D., & Papaoikonomou, E. (2020). *Sustainable development goals: A powerful framework for embedding ethics, CSR, and Sustainability in Management Education*. MDPI. https://www.mdpi.com/2071-1050/12/5/1762

Sulkowski, A. (2020). *Tao OF Dao: Hardcoding business ethics on blockchain*. https://www.researchgate.net/profile/Adam-Sulkowski-2/publication/336316096_THE_TAO_OF_DAO_HARDCODING_BUSINESS_ETHICS_ON_BLOCKCHAIN/links/5e8dd07ba6fdcca789fe0a34/THE-TAO-OF-DAO-HARDCODING-BUSINESS-ETHICS-ON-BLOCKCHAIN.pdf

Surthi, S. (2020). *New paradigms in business management practices*. https://www.amazon.in/NEW-PARADIGMS-BUSINESS-MANAGEMENT-PRACTICES-ebook/dp/B08RQR5Q9N/ref=sr_1_2?dchild=1&qid=1633022006&qsid=260-5567715-1581557&refinements=p_27%3ASruthi%2BS&s=digital-text&sr=1-2&sres=B07KKKK7NQ%2CB08RQR5Q9N%2CB08RP1MX5P%2CB08R7XQRFW%2CB07DPMZ1PZ&text=Sruthi%2BS

Tsai, H. J., & Wu, Y. (2021). Changes in Corporate Social Responsibility and Stock Performance. *Journal of Business Ethics*. Advance online publication. doi:10.100710551-021-04772-w

Weatherly, N. L. (2021). The ethics of organizational Behavior Management. *Journal of Organizational Behavior Management*, *41*(3), 197–214. doi:10.1080/01608061.2021.1890664

Whiting, R., & Pritchard, K. (2017, December 1). *Digital Ethics*. Birkbeck Institutional Research Online. https://eprints.bbk.ac.uk/id/eprint/15623/

Yang, J., & Basile, K. (2021, April 20). Communicating corporate social responsibility: External stakeholder involvement, productivity and firm performance. *Journal of Business Ethics*. https://link.springer.com/article/10.1007/s10551-021-04812-5

Zayyad, H. M. (n.d.). *Corporate Social Responsibility and patronage intentions: The mediating effect of Brand Credibility*. Taylor & Francis. https://www.tandfonline.com/doi/abs/10.1080/13527266.2020.1728565

Chapter 13
Fintech and Blockchain:
Maximizing Benefit and Minimizing Harm

Jane Thomason
Centre for Blockchain Technology, University College London, UK

ABSTRACT

The rapidly growing field of Blockchain and Decentralised Finance (DeFi) has the potential to transform many aspects of the financial world. It offers an abundance of opportunities to reduce costs, increase transparency and reduce the need for middlemen in financial services. While this promise of automation and decentralization is attractive, it is important to consider the potential for inadvertent or deliberate automation of unethical conduct at scale. Ethical questions involve the consideration of conflicting moral choices and dilemmas. Blockchain creates ethical dilemmas for developers, investors, consumers, and regulators at the technology, application, and societal levels. This chapter provides a perspective on the emerging field of DeFi and Blockchain in financial services, a reflection on the ethical questions that arise, how they are being addressed, the key issues, and further research needed in this growing field of Blockchain ethics. The objective of the chapter is for students, developers, CEO's and governments to appreciate the moral and ethical issues being uncovered in the course of the development and deployment of Blockchain technology. Blockchain, as with all technology, is a tool and is as beneficial and useful as the care that is taken to make it. There remains a need to ensure that Blockchains are built and deployed with due concern for ethics.

INTRODUCTION

Blockchains are characterized by (1) a transparent and publicly auditable database; (2) cryptographically secured data; (3) immutability, the record of the transaction in the database cannot be changed; (4) a decentralized network, thousands of computers validating each record of the ledger; and (5) timestamped transactions (Thomason et al., 2019, p. 25). The powerful combination of automation and Blockchain decentralization is reshaping the financial system in multiple ways, including the storage of digital records (identities, assets, voting rights, etc); the exchange of digital assets (via direct peer-to-peer transactions), and the recording and execution of smart contracts.

DOI: 10.4018/978-1-7998-8467-5.ch013

Blockchain has a key role in fintech innovation and has driven significant and wide-reaching innovation throughout the finance industry, from the latest innovations in crypto-assets to DeFi and Non Fungible Tokens (NFTs). To illustrate the rapid growth, Novum Insights (2021) reported that in 2021 crypto market capitalization peaked at $2 trillion, and the Total Value Locked (TVL) in DeFi was $54 billion in June 2021.

Despite being traditionally cautious, banks have been some of the earliest adopters of Blockchain technology, and over 200 banks (Joshua, 2020) and over 40 central banks are experimenting with Blockchain to varying degrees of depth, interest, and progress (World Economic Forum, 2019). Blockchain proof of concepts and deployments are underway for; the establishment of syndicated loan joint ventures (Keirns, 2017); provision of utility settlement coins[1]; settled high-value securities transactions; mutualized KYC servicing; efficient cross-border transaction and verification; internal foreign exchange balance sheet reconciliation (Gonçalves, 2019); back-office business management; secure interbank letters of credit; trade finance transactions; and bond issue and settlement.

The benefits of Blockchain for process and transaction management include; wide-ranging oversight of trades from trade to settlement, transparency and real-time access to a shared ledger, automation and reduced reliance on external settlement networks, efficiency gains, and reduced counterparty, market, and credit risk (Gonçalves, 2019; Costello, 2017). Accenture estimates that utilizing Blockchains could enable investment banks to save US$8 billion on a cost base of US$30 billion - a 27% cost-saving.

The implementation of institutional measures of transparency and automation are likely to reshape key financial, operational risk and finance systems from manual systems on cloud platforms to automated systems on shared data platforms. Currently, despite the importance of data reconciliation and accuracy, the majority of institutions maintain their own data, creating inefficiencies: duplication, input failures, repeated and ineffective consultation. Blockchain can transform today's multiple and sequential data reconciliation models to one where reconciliation is an integral part of the transactional process. However with the potent mix of automation and financial technologies powering faster transactions and accumulating masses of consumer data surfaces a plethora of ethical questions and demands new approaches and frameworks to ensure that ethics are baked into the design of these new and powerful technologies.

Decentralised Finance (DeFi)

DeFi enables cryptocurrency holders to make use of traditional financial services such as borrowing, lending, trading, and investing in a decentralized and transparent manner. DeFi is the opening of the traditional, closed, financial system, to a modular, interoperable, and programmable system. In 2020 and 2021, there has been spectacular growth of DeFi, opening a Pandora's Box of ethical challenges.

The governance of DeFi projects is a topic that requires new thinking and approaches. This is because DeFi projects commonly use Decentralised Autonomous Organisations (DAOs; Hsieh et al., 2018) to help move cryptocurrencies across different Blockchains, and for lending and yield farming. DAOs are "non-hierarchical organizations that perform and record routine tasks on a peer-to-peer, cryptographically secure, public network, and rely on the voluntary contributions of their internal stakeholders to operate, manage, and evolve the organization through a democratic consultation process" (Thomason, 2021). It is estimated that DAOs oversee more than $480 million (Beck & Asher, 2021).

DAOs are open-source and transparent (Voshmgir, 2019) but depending on the governance rules, have different levels of decentralization. While the network might be geographically decentralized, and with many independent but equal network actors, the governance rules written in the smart contract can be a point of centralization. There are various points at which DAOs can be decentralised, architecturally with independent actors running different nodes, geographically decentralized, but they are logically centralized through the protocol. DAOs also have both internal and external governance components (Zwitter & Hazenberg, 2020). Internal governance tends to be characterized by non-hierarchical modes of governance with quasi-democratic features, while the external governance is reliant on clusters of servers and individual nodes for the functioning of the network and decision-making. The reality is that those who control nodes and server capacity can exert more influence on decision-making.

The infamous DAO Hack demonstrated how vulnerable DAO governance can be. In this case, a smart contract both granted investors voting rights, as well as decisions regarding the distribution and management of the $150 million dollar fund. Voting was achieved through consensus of the investing community. However, there were no contingencies to define, manage, or control conflicts within the investing community. DAO governance operations were in the hands of an algorithm which was the DAO's sole governance mechanism. It operated as it was instructed and according to previously-agreed rules. The attackers cleverly exploited a vulnerability of The DAO's Blockchain-encoded smart contract (Zwitter & Hazenberg, 2020).

Morrison et al. (2020) summarise some of the key features of DAOs, which give rise to risks. These are (1) there are no trusted human executives as the entity is governed and operated by smart contracts, (2) the smart contracts are written and executed as computer code, (3) monitoring and enforcement of smart contracts are by computer algorithms, (4) there are weak or non-existent mechanisms for dispute resolution, as participants have agreed in advance to abide by the code of the smart contract(s). In contrast to conventional organizations, in a DAO, IT governance and corporate governance are controlled by code. This raises ethical questions about how details of governance, legalities, consumer protection and the logic flaws in the code are corrected and who has the liability for losses.

KEY ETHICAL ISSUES AND QUESTIONS

Blockchain ethics can be examined at three levels (Tang et al., 2019) - the technology stack, applications, and society.

Technology Stack

Proper conditions and methods of data sharing in Blockchain need to meet expectations in terms of security, privacy, efficiency, and system integrity. Blockchain establishes identities that are permanently linked to a unique individual and can be used in a variety of contexts to prove identity or credentials and move with the person. This requires multiple pieces of identifying information to create a digital identity.

Verification refers to ensuring the veracity of the information being entered into a Blockchain, and authentication refers to validating and accepting transactions on a Blockchain. Verification and authentication include questions such as who completes the verification and authentication and the methods by which this is done. For digital assets such as cryptocurrencies, the verification process is closely related to the transaction authentication process to determine if the entity initiating a transaction has control over that asset. However, when linking a non-digital asset like a person or an object to a Blockchain, verification is more complicated because it requires human interaction and the political, legal, and ethical interactions that come with it.

The definition, granting, and execution of access are critical to people's ability to use and interact with a Blockchain system. Access to individuals' personal information on a Blockchain may have serious consequences for individuals if that information is exploited. Access also includes more intangible questions around digital literacy and the effective ability to access the system.

The key ethical question at the technology stack level is "how should data security, privacy, and accessibility be ensured ethically, and what ethical information management strategy should be applied in system development and use?"

Applications

It is in the applications that are built on blockchains, that ethical questions emerge, as this rapidly growing industry expands and innovates with crypto, DeFi, Non Fungible Tokens (NFTs) and smart contracts.

Cryptocurrency risks include consumer protection, money laundering, criminal abuses, volatility and tax evasion, among others. At a society level, they have the potential to challenge the international monetary system and Bitcoin has been criticized for costly, energy hungry and otherwise meaningless computations used in coin mining. Advances in the underlying mechanisms are needed to make cryptocurrencies more ethical and sustainable.

DeFi has introduced an array of ethical challenges, many related to consumer protection. For example, the decentralized exchanges order book approach is susceptible to manipulation, an abusive trader can place large buy or sell orders to mislead the market sentiment and cancel the orders, creating high slippage for other traders. Automated Market Makers (AMM) face significant risks of impermanent loss, as the prices of the tokens held in a pool are determined by an algorithm that adjusts the ratios of the tokens in the pool, if the ratio between the two tokens changes drastically after depositing them in the pool, there will be a high slippage. Yield farming risks include liquidation risk, technical risk, and price risk. The value of the token as the collateral to take out a loan could drop below the price of the loan.

DeFi individual project governance can be opaque. DeFi platforms rely on open-source computer code, and some pay security researchers to conduct audits of the code to see if there are any vulnerabilities. Unaudited projects are riskier than audited, but just because a project has been audited does not mean that it is safe. If the smart contract malfunctions, is hacked, or otherwise has a problem, there is no recourse.

DeFi governance (Stroponiati et al., 2020) can also create an ethical minefield. Most DAOs raise money and in return, investors get back governance tokens, thus creating a high degree of centralization at the start of token distribution. As Protocols started using their governance tokens as "rewards'' for users participating in the network, many users see tokens as yield, not voting rights. There is usually no minimum number to initiate the governance but in order for a system to be considered sufficiently decentralized, there needs to be a high minimum number of token holders. Thus, the economic incentives of providing liquidity in order to get rewarded with governance tokens, encourages competitive and speculative behavior (Cousaert, 2021) which leads to a centralized governance structure, since tokens slowly concentrate in a few hands. Projects can also become vulnerable to attacks because of excessive centralization and parties with conflict of interest can push through proposals, and activist investors can acquire a significant number of governance tokens to help push through proposals profitable to them.

Smart contracts raise ethical questions about self-executing code that operates autonomously and raise questions of legal jurisdiction and issues of territoriality. When an organization is governed and operated by smart contracts, the smart contracts which form the governance are written and executed as computer code. The monitoring and enforcement of smart contracts are by computer algorithms, and there are often weak or non-existent mechanisms for dispute resolution, since all participants have agreed in advance to abide by the code of the smart contract. There is no legal recognition of documents or financial instruments stored on or issued for Blockchains. When a smart contract fails, under which law and in which jurisdiction can action be taken? The legal status of a DAO is also a gray area (Okaformbah, 2019), as nobody owns the organization, who can be sued and who sues or in the case of liquidating a tangible asset owned by the DAO, what rules are to be followed?

In this experimental phase of a new technology should consumers and investors be protected? when activities are not controlled, moderated, intermediated, hosted, or validated by a single or center and how can regulators protect consumers and investors when there are not any intermediaries to regulate, as it's totally P2P? Should the people who write the smart contracts be held accountable for the smart contracts?

Institutions and Society

In order for digital assets to be widely distributed in any given ecosystem, there will be some degree of adaptation with traditional ways of operating by introducing regulation. Broader frameworks are also required for recognizing Blockchain records, determining the legal status of tokens, and harmonizing the relationship between the General Data Protection Regulation (GDPR) right to be forgotten and the immutable nature of blocks.

Other challenges for regulators include decentralizing the financial system, managing economic stability, and protecting consumer interests. It is traditional for central banks to control currency issuance to tightly protect national monetary sovereignty. Cryptocurrencies are in direct conflict with the established monetary systems and inevitably create ethical challenges for monetary policy (Dierksmeier & Seele, 2016).

Blockchain, like all technologies, can be misused, especially when there are risks of authoritarian states, persecution, and unintended consequences. There are risks of bad actors using digital identities, bank accounts, and mobile phones that allow authorities to track people's choices. Such control might allow authorities to increase surveillance over vulnerable or persecuted populations. An authoritarian state could use such data collected from refugees against refugees, or nations of the global North that have no sympathy for the movements of refugees and immigrants, to keep refugees in neighboring countries.

Blockchain economies demand a rethinking of governance. Blockchains make it possible to create leaderless, decentralized organizations which blur the jurisdictional boundaries of economic activity. In particular, the enforcement of accountability through technical specifications and smart contracts will require a deep understanding of the objectives of the network and decision rights, incentives, and accountabilities. Who should regulate digital, borderless economies? Blockchain decentralization at the social level facilitates the shift from centralized human governance to decentralized algorithm governance. Who should be responsible?

Blockchain has the potential to introduce disruptive forms of innovation that take organizations forward into a new era of connected digitization. At a societal level, where a new technology is highly disruptive, should there be a balance between the interests of the sovereign state and crypto builders and users and what ethical codes and regulations are needed for a future decentralized digital economy?

Issues, Controversies, Problems

Some commentators, such as Lindmark, argue that Blockchain technology deserves its own field of ethics (Orcutt, 2019). Others, such as Lapointe and Fishbane (2019), have developed tools to help identify and address ethical risks. Their Blockchain Ethical Design Framework (Lapointe & Fishbane, 2019) outlines six issues for ethical consideration: governance, identity, verification and authentication, access, ownership of data, and security. They identify guiding questions to identify the effects of the design choices on the end-users and communities: (i) How is governance created and maintained? (ii) How is identity defined and established? (iii) How are inputs verified and transactions authenticated? (iv) How is access defined, granted, and executed? (v) How is ownership of data defined, granted, and executed; and (vi) How is security set up and ensured?

Data Ethics

Underpinning all of the discussion about Blockchain and DeFi ethics is data. Floridi and Taddeo (2016) describe data ethics as:

a new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing, and use), algorithms (including artificial intelligence, artificial agents, machine learning, and robots), and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g., right conducts or right values; p. 3)

Data ethics shift the level of abstraction of ethical inquiries from being information-centric to being data-centric. Ethical analyses need to concentrate on the content and nature of computational operations and the interactions among hardware, software, and data, including data that never translates directly into information but can be used to support actions.

Data ethics studies evaluate moral problems related to data (e.g., including generation, recording, curation, processing, dissemination, sharing, and use), algorithms (e.g., artificial intelligence, artificial agents, machine learning, and robots), and corresponding practices (e.g., responsible innovation, programming, hacking, and professional codes), to form and support morally good solutions (e.g., right conducts or values). Key ethical issues include the re-identification of individuals through data-mining, linking, merging, and re-using large datasets, as well as risks for group privacy, when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination to group-targeted forms of violence (Floridi & Taddeo, 2016).

Automation has significantly increased processing speed but also led to more ethical risks and the need to examine unintended consequences of an automated technology. Scott (2018) raises some powerful ethical questions related to automation, including what happens when there is a mistake or vulnerability in the code?

Data ethics should be a consideration in every project - how will the ethics of data, algorithms, and practice be considered during design. How will data be governed, informed consent, customer understanding, the requirement for all data to be necessary for the purpose for which it is collected, and the extent of human oversight or intervention required in the decision process. A highly proximate relationship strengthens accountability when things go wrong. However, there is a risk when leaders lack visibility into the creation and deployment of algorithms they may not apply sufficient controls. Lee explains, "proximity denotes both how physically close or emotionally close we are to someone or something" (Lopez, 2019, para. 24). "Our world is increasingly distant and non-proximate in nature, resulting in our leaders increasingly using amoral, cost-benefit analysis when making decisions" (Lopez, 2019, para. 25). Thus, the closer one is to a problem space, the more ownership or accountability they have.

Data privacy in regard to Blockchains is a complex issue that arises mainly due to the immutability of data on the Blockchain. The situation is especially complex for personal data related to an identified or identifiable person. The European Union's General Data Privacy Regulation (GDPR) (2018), provides a set of regulations to ensure that the EU can guarantee the protection of individual data. The GDPR, however, was written with a centralized entity in mind that has the power to control access rights, which is not the case when Blockchain technology is used. It is thus unclear how Blockchain technology will compete with the GDPR (Posadas, 2018).

Consumer Vulnerability

At the heart of digital ethical considerations for Blockchain lies consumer protection. The vast amount of personal and private data stored creates a number of points for consumer vulnerability. For example, to enable the distribution of cash-for-food aid, the World Food Programme's (WFP) Building Blocks initiative collects personal data, including biometrics, from over 500,000 Syrian refugees in Jordan (Rugeviciute & Mehrpouya, 2019). Personal data, entitlements, and transaction logs are stored on the Ethereum Blockchain to provide a virtual bank account and ID for each refugee. While the Building Blocks platform has been successful, there is a risk that conducting iris scans on refugees in shops robs them of dignity. Sensitive, personally identifiable information for some of the most vulnerable people in the world is also being generated and made accessible across agencies, inevitably introducing a greater risk of data breaches. Refugees and vulnerable people might give up personal (including biometric) information about themselves, stored on an immutable ledger, in return for temporary support with basic necessities. These data could also be used in the future to make decisions about individuals with far-reaching consequences. Some may suffer punitive restrictions based on decisions made using biased algorithms calculating, for example, the risk of absconding or working without a permit. "Consent has three parts to it. One of them is, I must be adequately informed. Two, I must be able to make a decision. And three, I need to be able to understand and appreciate the information to make that decision" (Van Leeuwen, 2020, p. 4). Data shared by some of the most persecuted on a Blockchain highlights the obvious need for a more robust regulatory framework to effectively mitigate the risks associated with data protection, privacy, and human rights.

Decentralized Autonomous Organizations

Morrison et al. (2020) describe how DAOs operate: 1) there are no trusted human executives since the organization is governed and operated by smart contracts. 2) DAO governance is provided by smart contracts that are written and executed as computer code, and 3) monitoring and enforcement of smart contracts are also run by computer algorithms. Finally, there can be weak or non-existent mechanisms for dispute resolution since all participants have agreed in advance to abide by the code of the smart contracts. While DAOs have been widely promoted by Blockchain proponents as providing transparency and trust, it is increasingly evident that DAOs are still in an experimental stage; there are also a large number of information asymmetries that may exist in a DAO and participant ambitions, motivations, values, or priorities are not transparent. Where priorities and values do not align and there are no contingencies to define, manage, or control these conflicts, DAOs may not provide adequate governance to protect consumers.

Regulation

Regulatory issues are particularly formative in the case of Blockchain. There is no generally applicable mechanism for adjudicating disputes arising from transactions that are executed with digital assets. When automatically executable contracts such as those that underpinned the DAO are exploited, there is little legal recourse for those affected. Although "certain operational clauses in legal contracts" may be automated to beneficial effect (International Swaps and Derivatives Association [ISDA], 2017), the principle of "code is law" may not be workable without a suitable legal framework. Goodell and Aste (2019) make the point that for Blockchain to be widely distributed in any given investment community, there must be some degree of adaptation with traditional ways of operating by introducing regulation. This means Anti-Money Laundering Regulations/Know Your Customer (AML/KYC) processes that provide guidelines for financial institutions to comply with to prevent fraud. Broader frameworks are also required for such purposes, such as recognizing Blockchain records, determining the legal status of tokens, and harmonizing the relationship between the General Data Protection Regulation (GDPR) 'right to be forgotten and the immutable nature of blocks.

In summary, a plethora of issues, controversies and problems remain to be solved, as this rapidly developing technology evolves. Data ethics and consumer protection will be a key focus for ethicists and regulators alike. However, the technology is developing faster than many can keep up with, so a collaborative approach will be needed to ensure that Blockchain ethics are at the heart of all technology designs.

SOLUTIONS AND RECOMMENDATIONS

Hyrynsalmi et al. (2020) propose that stakeholders need to understand the ethical and moral advantages and challenges of Blockchain, that Blockchain's technology stack needs to incorporate core ethical values, that the applications and business models built with Blockchain should respect those ethical values, and that regulations imposed on Blockchain applications must encode ethical principles. The Kalifa Report (Kalifa, 2021) recommends that the government should consider undertaking a review of the future legal and regulatory framework for the role of ethics in AI models (e.g., should there be regulation of models that might satisfy the requirement for fairness but nevertheless could lead to bias or discrimination?). Things will only get more complicated in the future.

Where Blockchain is being deployed as part of an organisation, it is relatively straightforward. In the end, management is responsible when things go wrong. The people behind the system must be truly responsible and ensure that the system protects customer data. Projects which can demonstrate a real commitment to ethical practice and leadership will have a competitive advantage (Makepeace, n.d.). Linklaters (2019) on ethics in banking and finance highlight three major aspects for ensuring ethical practices. The first is the role of the board in articulating a culture of risk awareness and ethical behavior, and the second is how leadership and staff at all levels can shape the culture of an organization. The third way is how the risk management and control framework embed and monitor ethical values within an organization.

However, when dealing with DAOs, where all governance is automated and decentralized, this becomes more complicated. The ethical issues need to be considered during design and built into the code and algorithms and execution plans. This is not impossible, but the majority of software developers have not had training in ethics, how to think about ethical questions and answering them. At this stage, there is not a professional association for software developers, with a codified ethical standard. There do exist, however, a number of pledges and codes of conduct, such as the ACM Code of Ethics (ACM Code 2018 Task Force, 2018).

Singapore has long been a trendsetter in terms of digital innovation, and this remains the case in considering digital ethics in Blockchain. The Monetary Authority of Singapore, in conjunction with domestic financial institutions and Microsoft and Amazon Web Services, launched its fairness, ethics, accountability, and transparency (Feat) principles (Mackhight, 2019) for the use of AI and data analytics in decision making in 2018. They sought to develop this as a co-creation with industry. Singapore has taken a high-level, principles-based approach because it was targeting the whole financial ecosystem, regulated and non-regulated entities, from the smallest fintech startups to the large banks and tech giants.

This is a rapidly developing field, but there is certainly scope for greater ethical reflection and purposeful consideration of ethics in the design of Blockchain and DeFi applications. This should include:

1. Encourage greater ethical reflection from developers during design.
2. Connect developers more closely to the ethical outcomes of their decisions and algorithms.
3. Encourage community and network to take a more active and demanding stance on ethics.
4. Encourage community and network to understand what is happening behind the scenes with governance and decisions.

FUTURE RESEARCH DIRECTIONS

Interdisciplinary research involving ethicists, philosophers, computer scientists, economists, political scientists, law experts, sociologists, psychologists, management scientists, and anthropologists, who can identify the potential implications of Blockchain far beyond technology is needed. A systematic, cohesive, and joint research agenda informed by stakeholders' views and roles in conceptualizing, developing, and delivering Blockchain technologies will be important. There is a need for research at the intersection of financial technologies and their adoption, with reference to building capabilities in dealing with externalities such as legal and compliance issues.

Further research topics include:

- Further research into existing codes, and professional organisations for coders, and identifying the pathway for the establishment of a globally adopted ethical code of conduct for developers, would be a valuable step.
- The development of new rapid-cycle analytics research techniques to keep up with the pace of implementation of Blockchain projects and monitor their impact.
- New approaches to data ethics delineated by three axes of research, including the ethics of data, algorithms, and practice.

- How and at which level (local, regional, national, global) of governance can the negative externalities of Blockchain technology best be regulated?
- To identify and quantify the effects of digital assets on the effectiveness of tools used by the central banks.
- When is ID tracking ethically advantageous and in which scenarios does it create a moral hazard?
- Research into regulatory questions of legal jurisdiction and issues of territoriality, legal recognition of documents or financial instruments stored on or issued on Blockchains.
- How with decentralization of the financial system can regulators manage economic stability and protect consumer interests?
- Appropriate regulatory measures related to equity requirements, safe experimentation of new technologies, and less cumbersome supervisory arrangements?

CONCLUSION

New and powerful technologies demand a new approach and one that is co-developed with industry. Further research is needed into the development of and better use of ethical frameworks and criteria to ensure technology is building out in an inclusive, systemic way to address the issues it is supposed to solve. Ethics enables us to make judgments about what should happen. Of all the ways you might act, which is the best? There is a paucity of research into Blockchain Ethics and this is an emerging field.

Ethics raises more questions than answers. Even though the potential for Blockchain to transform many aspects of the world is there, there remains a need to ensure that the technology is built and deployed with due concern for ethics. Blockchain can have ethical impacts at the technology, application, and societal levels. It is important that these are considered and built into system design with intentionality. While the promise of automation and decentralization is attractive, it is important to avoid the inadvertent facilitation of unethical conduct. Blockchain technology is a conditional good; it is only as beneficial and useful as the care that is taken to make it.

REFERENCES

ACM Code 2018 Task Force. (2018). *ACM Code of ethics and professional conduct*. Association for Computing Machinery Committee on Professional Ethics. https://www.acm.org/code-of-ethics

Beck, J., & Asher, M. (2021, February 3). Why Decentralized Autonomous Organizations (DAOs) are essential to DeFi. *Consensys*. https://consensys.net/blog/codefi/daos/

Costello, E. (2017, August 31). *Big banks back Blockchain for back-office business*. International Investment. https://www.internationalinvestment.net/internationalinvestment/news/3504643/banks-blockchain-office-business

Cousaert, S. (2021, March 25). *Generalizing knowledge on DEXs with AMMs — Part I*. Medium. https://medium.com/uclcbt/generalizing-knowledge-on-dexs-with-amms-2963d07ebac7

Dierksmeier, C., & Seele, P. (2016). Cryptocurrencies and business ethics. *Journal of Business Ethics*, *152*(1), 1–14. doi:10.100710551-016-3298-0 PMID:30930508

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions - Royal Society. Mathematical, Physical, and Engineering Sciences*, *374*(2083), 20160360. Advance online publication. doi:10.1098/rsta.2016.0360 PMID:28336805

Gonçalves, P. (2019, January 15). *HSBC banks on Blockchain tech to process $250bn worth of FX transactions*. International Investment. https://www.internationalinvestment.net/news/4000471/hsbc-banks-blockchain-tech-process-usd250bn-worth-transactions

Goodell, G., & Aste, T. (2019). Can cryptocurrencies preserve privacy and comply with regulations? *Frontiers in Blockchain*, *2*, 4. Advance online publication. doi:10.3389/fbloc.2019.00004

Hsieh, Y.-Y., Vergne, J.-P., Anderson, P., Lakhani, K., & Reitzig, M. (2018). Bitcoin and the rise of decentralized autonomous organizations. *Journal of Organization Design*, *7*(14), 14. Advance online publication. doi:10.118641469-018-0038-1

Hyrynsalmi, S., Hyrynsalmi, S. M., & Kimppa, K. K. (2020). Blockchain ethics: A systematic literature review of Blockchain research. In M. Cacace, R. Halonen, H. Li, T. P. Orrensalo, C. Li, G. Widén, & R. Suomi (Eds.), *Well-Being in the information society. Fruits of respect* (Vol. 1270, pp. 145–155). Springer. doi:10.1007/978-3-030-57847-3_10

International Swaps and Derivatives Association. (2017). *Smart contracts and distributed ledger - A legal perspective* [White paper]. https://www.isda.org/a/6EKDE/smart-contracts-and-distributed-ledger-a-legal-perspective.pdf

Joshua. (2020, January 11). *Comprehensive list of banks using Blockchain technology in 2020*. Hackernoon. https://hackernoon.com/comprehensive-list-of-banks-using-blockchain-technology-in-2020-revised-and-updated-uq493yrb

Kalifa, R. (2021). *Kalifa review of UK Fintech*. UK Treasury. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/978396/KalifaReviewofUKFintech01.pdf

Keirns, G. (2017, March 30). *Major banks, start-ups advance syndicated loan pilots*. Coindesk. https://www.coindesk.com/banks-startups-Blockchain-syndicated-loans

Lapointe, C., & Fishbane, L. (2019). The Blockchain ethical design framework. *Innovations: Technology, Governance, Globalization*, *12*(3–4), 50–71. doi:10.1162/inov_a_00275

Linklaters. (2019). *Leadership, governance, systems and controls*. https://www.linklaters.com/en/insights/publications/2019/february/ethics-in-banking-and-finance/leadership-governance-systems-and-controls

Lopez, P. (2019, July 23). *Five principles of ethics to help drive Fintech platform risk management*. Linkedin. https://www.linkedin.com/pulse/five-principles-ethics-help-drive-fintech-platform-risk-lopez/

Mackhight, J. (2019, April 1). *An ethical framework for the AI age*. The Banker. https://www.thebanker.com/Transactions-Technology/An-ethical-framework-for-the-AI-age

Makepeace, N. (n.d.). Data ethics starts at the top. *GWI*. https://gwi.com.au/blog/data-ethics-starts-at-the-top/

Morrison, R., Mazey, N. C., & Wingreen, S. C. (2020). The DAO controversy: The case for a new species of corporate governance? *Frontiers in Blockchain*, *3*, 25. Advance online publication. doi:10.3389/fbloc.2020.00025

Novum Insights. (2021, June 30). *Macro overview of the crypto economy 2021 | Upcoming Novum Insights DeFi report 2021*. https://novuminsights.com/post/macro-overview-of-the-crypto-economy-2021-or-upcoming-novum-insights-defi-report-2021/

Okaformbah, C. (2019, February 19). *Governance in a Decentralized Autonomous Organization*. Medium. https://justcharles.medium.com/governance-in-a-decentralized-autonomous-organization-425f56b3e8bb

Orcutt, M. (2019, October 10). Why it's time to start talking about Blockchain ethics. *MIT Technology Review*. https://www.technologyreview.com/2019/10/10/132652/why-its-time-to-start-talking-about-Blockchain-ethics/

Posadas, D. V., Jr. (2018). The internet of things: The GDPR and the Blockchain may be incompatible. *Journal of Internet Law, 21*(11), 1, 20–29.

Rugeviciute, A., & Mehrpouya, A. (2019). Blockchain, a panacea for development accountability? A study of the barriers and enablers for Blockchain's adoption by development aid organizations. *Frontiers in Blockchain*, *2*, 15. Advance online publication. doi:10.3389/fbloc.2019.00015

Scott, B. (2018). Hard coding ethics into Fintech. *Finance & the Common Good / Bien Commun*, 44–45, 80–93.

Stroponiati, K., Abugov, I., Varelas, Y., Stroponiatis, K., Jurgeleviciene, M., & Rao, Y. S. R. (2020). Decentralized governance in DeFi: Examples and pitfalls. *Squarespace.* https://static1.squarespace.com/static/5966eb2ff7e0ab3d29b6b55d/t/5f989987fc086a1d8482ae70/1603837124500/defi_governance_paper.pdf

Tang, Y., Xiong, J., Becerril-Arreola, R., & Iyer, L. (2019). Ethics of Blockchain: A framework of technology, applications, impacts, and research directions. *Information Technology & People*, *33*(2), 602–632. doi:10.1108/ITP-10-2018-0491

Thomason, J. (2021, May 23). *Why DAO Governance Matters for DeFi* [Article]. LinkedIn. https://www.linkedin.com/pulse/why-dao-governance-matters-defi-dr-jane-thomason

Thomason, J., Bernhardt, S., Kansara, T., & Cooper, N. (2019). *Blockchain technology for global social change*. IGI Global. doi:10.4018/978-1-5225-9578-6

Van Leeuwen, D. (2020). *Bioethics: Autonomy. For me, on behalf of me: An interview with Kenneth Goodman*. Health Hats. https://www.health-hats.com/wp-content/uploads/2020/02/20200209_HHP059_Kenneth_Goodman_FINAL.pdf

Voshmgir, S. (2019, July). Tokenized networks: What is a DAO? *BlockchainHub*. https://blockchainhub.net/dao-decentralized-autonomous-organization/

World Economic Forum. (2019, March). *Central banks and distributed Ledger technology: How are central banks exploring Blockchain today?* [Whitepaper]. https://www3.weforum.org/docs/WEF_Central_Bank_Activity_in_Blockchain_DLT.pdf

Zwitter, A., & Hazenberg, J. (2020). Decentralized network governance: Blockchain technology and the future of regulation. *Frontiers in Blockchain*, *3*, 12. Advance online publication. doi:10.3389/fbloc.2020.00012

ADDITIONAL READING

Armer, P. (1968). *Privacy aspects of the cashless and checkless society* [Testimony before the Senate Subcommittee on Administrative Practice and Procedure]. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/papers/2013/P3822.pdf

Australian Law Reform Commission. (2014, March 30). *Serious invasions of privacy in the digital era (DP 80)*. Australian Government. https://www.alrc.gov.au/publications/serious-invasions-privacy-dp-80

Beckett, L. (2014, June 13). Everything we know about what data brokers know about you. *ProPublica*. https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you

Christl, W. (2017). *Corporate surveillance in everyday life: How companies collect, combine, analyze, trade, and use personal data on billions*. Cracked Labs. https://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf

Competitive Enterprise Institute. (2000). *The future of financial privacy*. https://cei.org/studies-books/future-financial-privacy

Falkon, S. (2017, December 24). *The story of the DAO – Its history and consequences*. Medium. https://medium.com/swlh/the-story-of-the-dao-its-history-and-consequences-71e6a8a551ee

Goodell, G., & Aste, T. (2018, April 17). *Blockchain technology for the public good: Design constraints in a human rights context*. Open Access Government. https://www.openaccessgovernment.org/Blockchain-technology-for-the-public-good-design-constraints-in-a-human-rights-context/44595/

Maxwell, W., & Salmon, J. (2017). *A guide to Blockchain and data protection*. Hogan Lovells. https://www.hlengage.com/_uploads/downloads/5425GuidetoBlockchainV9FORWEB.pdf

Tasca, P., & Aste, T. (2018, July 18). *Crypto assets and the regulator's role: Ignore, regulate or kill?* Open Access Government. https://www.openaccessgovernment.org/crypto-assets-and-the-regulators-role-ignore-regulate-or-kill/47858/

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75–89. doi:10.1057/jit.2015.5

KEY TERMS AND DEFINITIONS

Algorithm: A process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

Blockchain: One form of distributed ledger technology. A peer-to-peer method of secure data transmission using grouped "blocks" of encrypted data.

Cloud Computing: Cloud computing is the delivery of different services through the Internet, including data storage, servers, databases, networking, and software.

Consensus Algorithm: A process used in computer science to achieve agreement on a single data value across distributed networks. A consensus algorithm is designed to solve a consensus problem to achieve network reliability across multiple nodes.

Defi: The opening of the traditional, closed, financial system, to one that is modular, interoperable, and programmable.

Distributed Autonomous Organisation: Are non-hierarchical organizations that perform and record routine tasks on a peer-to-peer, cryptographically secure, public network, and rely on the voluntary contributions of their internal stakeholders to operate, manage, and evolve the organization through a democratic consultation process.

Know Your Customer (KYC): This process refers to a project's or financial institution's obligations to verify the identity of a customer in line with global anti-money laundering laws.

ENDNOTE

[1] Utility settlement coins provide a digital cash instrument, the equivalent of a central bank-backed currency, to introduce efficiencies into financial market clearing and settlement. Fnality (proposed for a 2020 launch) would issue a coin convertible at parity, backed by a central bank's currency. CAD, EUR, GBP, JPY and USD are supported.

Chapter 14
Ethical Risks in the Cross Section of Extended Reality (XR), Geographic Information Systems (GIS), and Artificial Intelligence (AI)

Monika Manolova
MKAI, Bulgaria

ABSTRACT

The ethical risks which emerge from the cross section of artificial intelligence, extended reality, and geographic information systems could be examined in two broad categories of environmental and user-centric interactions of human beings with AI-curated mixed realities. These categories resonate with the capacity of AI to significantly impact the efficient application of extended reality technologies, while utilizing geodata and behavioral modelling to alter and transform experiences. While regulatory frameworks are catching up with the rights of users in the digital economy, the recently accelerated growth of immersive technologies provides further scenarios and use cases, which ought to be considered for their capacity to amplify biases, produce alternative realities, and affect human emotions.

INTRODUCTION

Within 2020 the cross section of data science and geographic information systems (GIS) expanded beyond the textbooks and into the daily consumption of internet audiences with the global consumption of the Johns Hopkins COVID-19 dashboard (Johns Hopkins University, 2021). It was a daily visited page as we followed the progress of a pandemic across the world. Data science stretched beyond its statistics roots into an explainable, shareable visual of events on scale outside human comprehension.

The mathematical ramifications obvious for the public as figures became a base for policies and solutions with significant effect on the daily functioning of societies, while means and trends had a

DOI: 10.4018/978-1-7998-8467-5.ch014

demonstrated reflection in the way societies operate (Committee for the Coordination of Statistical Activities [CCSA], 2020).

The mainstreaming of the data science field was accompanied by expanded daily hours in front of screens and emergence of a brand-new internet minute, which contains within itself not only an addicting human-digital symbiosis, but fundamental dependence and emotional investment into digital solution as a bridge between families, friends, communities, and societies.

Social geography classifies the waves of innovation diffusion into neat categories, which reach certain geographies within various times in history and provide for significant changes in perception, economic segmentation, policies, and societal norms (Sirk, 2020). But both national states and global powers such as the European Union (EU) and the United Nations (UN) have recognized the significance of the current stage of data fueling artificial intelligence solutions as unrivalled before in human history.

Artificial intelligence is the single unifying term used to describe a set of solutions that compile geometric progression of gathered data streams and provide quick, cheap solutions excluding hours upon hours of processing and human interference.

Regulations such as the General Data Protection Regulation in the European Union (EU) and the Data Act (European Commission, n.d.), and Electronic Communication Privacy Act, Cyber Intelligence Sharing and Protection Act (CISPA), The Health Insurance Portability and Accountability Act and the California Consumer Privacy Act (CCPA) within the United States represent the recognition of basic human rights within the wave of the new highly digitalized world. In 2021 the European Union proposed frontier regulation in artificial intelligence, which is positioned to be the first of its kind guardrail of ethical AI application in a wide set of use scenarios and a preventive measure for single entity utilization of advanced technologies in unlawful and harmful manner (Publications Office of the EU, 2020). The suggested measures further demonstrate the envisaged importance of artificial intelligence as a driver of economic, social, and political processes on global scale.

While the emergence of digital regulation is reassuring, the generational gaps in digital product consumption are constantly growing in sync with emerging technologies as the ways in which society consumes information (streaming, social networks, digital realities, immersive environment) and participates in economic exchanges (freelancer, influencer economics) continue to evolve along with cutting edge technologies, which were not available even five years ago.

In the post-pandemic disconnected physical world, social networks are accelerating their research in the provision a 3D experience instead of the php, html and Haskell-based, 2D experiences users have had thus far (Feldman, 2020), making platforms more immersive with encoded interactions and the capacity to gameplay inside a favorite TV programme and effect outcomes or follow favorite characters through storylines in 3D scenes. The addictive nature of the internet word transformed the internet minute into the first generation of internet lifetimes, where significant milestones in the human experience could exist almost entirely online.

The introduction of extended reality technologies further complicates the matter. The potential applications are endless and in a physically disconnected world, they could easily become addicting, the immersive nature of these environments becoming preferable to real life experiences. Large demographics already buy in the notions of reality TV, steaming entertainment, social media as forms of escapism and the lines between the real world and the digital one is becoming blurrier with the convergence between social networks and geo locations.

To add another layer of complexity, there are multiple geographies, cultures, religions, and belief systems, which are transformed by waves of technological evolution, by solutions beyond basic comprehension – brain computing, neural networks, linguistic processing, deep learning, and machine processing – the tales of science fiction.

COVID-19 is far from the last challenge humanity is facing. Climate change, water shortages, global hunger and physical ailments are some of the steepest climbs to move past, to be worthy of the expectation of the forthcoming generations as exemplified by UN's 17 millennium development goals (United Nations, n.d.). It is the children of the future that will experience this world in a more advanced way and will expect from the current generation to be the engineers behind the edge technologies that will make their world the kind of place, where no one is left behind.

I am speaking of the life of a man who knows that the world is not given by his fathers, but borrowed from his children; who has undertaken to cherish it and do it no damage, not because he is duty-bound, but because he loves the world and loves his children…" (Berry, 1971, p. 33)

Within that context, the recognition that regulations do not capture all dimensions of human-machine interaction and data consumption on all sides of the marketplace is fundamental for the establishment of sustainable ethical standards in line with the necessary accountability for the benefit of future generations.

ARTIFICIAL INTELLIGENCE(AI), EXTENDED REALITIES(XR) AND GEOGRAPHIC INFORMATION SYSTEMS (GIS)

Artificial Intelligence (AI)

While the concept of artificial intelligence has for long dominated the realm of science fiction, in the 21st century humanity is beginning to slowly unravel its multitude of applications to solve the mysteries of the universe or the mysteries of the human condition.

Artificial intelligence is an umbrella term for the advanced technologies, which mimic human cognition processes within computer systems ("Artificial intelligence," 2021). For data scientists AI toolboxes translate to sets of solutions for data processing via pipelines to automate conclusions. For GIS experts AI represents the computer vision to interpret heavyweight satellite imagery. For extended reality (XR) designers AI is a set of scenes aimed to further immerse a user within a specific sense or experience.

There are two widely accepted classifications of artificial intelligence (AI; Overby, 2020) – based on capacity and based on functionality[1].

The ethical issue which could arise of these classifications could be separated in two broad categories, which resonate with the ethical risks and dilemmas within the XR field:

1. Environment factors - focusing on the effects the environment has on a user.
2. Individual (user centric) factors – focusing on the effects various groups of users have on each other.

The external environment factors are related to national and global policies, vestiges of power and influence over the new data currency. Data mining, management, and the AI systems behind it is a growing economic frontier and it is bound to rearrange the global map.

External environment issues represent the interests of nations in the harvesting of the most advanced computational power, minds, and resources to utilize the economic prowess of data. Our global interconnected societies will continue to mount datasets that could contribute to improved environmental management, production cycles, healthcare, and education solutions. But those datasets could also be utilized for political and economic advantages, to lock certain groups and nations out of the global market, to draw further lines of separation based on nationalistic agendas.

Individual factors seem insignificant at first glance but provide no less food for thought and could affect society in no lesser manner. The individual users and their interactions with specific technologies will be the basis of further dilemmas. As society strives to reach the point of a super AI ecosystem, there will be a variety of scenarios that have not even been considered in the interactions of users and their assistant AIs, healthcare AI, even life partner AIs.

Regardless of whether it relates to machine learning, deep learning, natural language processing, computer vision, whether it is humanoid or not, artificial intelligence is the next step in the evolution of the data ecosystems utilized day-to-day.

Quicker, cheaper with a set of big queries or loops that improve upon the processes which are applied to daily functions remains the perfect business proposition, the single sentence pitch of a winning proposal, leading advancements, and economic prosperity. The solution businesses strive to achieve within the clear margins of numbers and figures. But where do human beings land in that optimized system?

The GAFAM five (Google (Alphabet), Apple, Facebook, Amazon, and Microsoft) have invested considerable funds in the development of first generation XR solutions, which employ artificial intelligence in cohesion with oculars to enhance experiences and to transition extended reality technologies beyond the realm of gaming and into the realms of communications, emergency services and education.

As AI will evolve to seamlessly process and fine tune our social network interaction so will extended reality (XR) evolve to produce an even more immersive online experience.

Extended Realities (XR)

Extended reality is a term which encompasses computer generated reality - virtual reality (completely computer simulated), mixed (hybrid between real and virtual world) and augmented reality (computer generated graphical overlay on top of the real world; Scribani, 2019).

While its conception relates to the world of entertainment in the form of computer games and immersive online experiences, it is expanding further into everyday exchanges in several imperceptible ways, which resemble the growth of the mobile phone industry from a luxury item to a daily appliance.

Ocular devices are becoming easier to acquire and a more acceptable item for users of all ages[2] and geographies, as we reside in a changing world with a global pandemic and a remote work phenomenon, social media engineering, society is changing along with its hardware and experiences, which may have seemed based solely in the domain of entertainment before 2020, are becoming far more attractive.

There are advantages to these advancements. The entertainment industry is adopting headsets as an extension of computer games (Unity already supplies vision immersion; Bardi, 2018) and improving experiences for a set of billion-dollar industries (including traditional cinemas). Outside of the realm of entertainment, these advancements would enhance building and construction, communications, healthcare, and emergency response training.

These applications would decrease costs in the modelling of smart city planning, provide a higher quality of training for emergency personnel and enhance education experiences – history, chemistry, philosophy, and travel experiences as never before in 3D in the center of the living room.

As with any innovation, there are ethical and design risks with XR as well. There are physical risks including prolonged effect of ocular devices on gaze and motor skills, there are data risks with the gathering of object-specific metadata. There are bias and behavioral risks in the sensemaking aspect of extended reality. Additional risks arise in the cross section of XR, geographic information systems and AI, which could be even more difficult to capture under the existing regulatory frameworks. While the EU Data Act envisages unbundling of some services and the proposed artificial intelligence regulations poses basic rights in the AI ecosystem, it is unlikely those will be able to regulate upcoming ethical risks as the utility of headsets lies predominantly in the provision of an expansive data flow from the user to the service.

Geographic Information Science (GIS)

While geographic information science (GIS) is a traditionally strategic scientific field employing long term modelling, the pace of innovation within global positioning systems (GPS) and Lidar technologies have placed GIS among one of the foremost data science niches. After all the data which is being computed and included within even the most simplistic of data science models is related to occurrences in the real world (geolocated in relation to individuals as exemplified by the Johns Hopkins dashboard on the intro section).

"Thus, geovisualization is conceived as a process rather than a product, although the term is also commonly used to refer to any visual display that features geospatial information (maps, images, 3D models, etc.). In the geovisualization process, the emphasis is on information exploration and sensemaking, where scientists and other experts design and use "visual geospatial displays to explore data, and through that exploration to generate hypotheses, develop problem solutions and construct knowledge." (Çöltekin et al., 2020, p. 231)

The second decade of the 21st century is transforming GIS into an expansive new science, which is oriented to incorporate various aspects of AI – including computer vision and smart mobility applications and to further incorporate the sensemaking aspect of XR within a smart city.

A separate branch of geographic information science exists due to its convergence with extended reality solutions XRGIS[3], which is concerned with the design of convincing 3D replicas and visualization as an extension of real-world objects (Arisona et al., 2020). The most recent example of these XR and GIS cross sections was the boom of PokemonGo in 2016. Geolocation is used daily by millions of users for directions and tagging, and an enhanced internet experience.

The advances of internet of things (IoT), machine learning, XR and AI solutions have allowed that branch of GIS to compute additional subliminal user metrics (gaze capture, head rotation, nonverbal cues) and to improve the user experience with additional sensemaking suggestions.

There are a multitude of applications of cross-sections of GIS, XR and AI that could predict the behaviors of users within a smart city network in the same way in which graph theory transport modelling is developed to predict the behaviors of autonomous vehicles within a certain network. A route to work calculated with the Djikstra algorithm carrying the possibility to be weighted with commercial preferences and prior visit history, with preferred services and AI curated content ("AI-Driven platform for content curation — or a Netflix of knowledge!," 2017).

Predictive modelling of the behaviors of subjects in their interactions with their location is a sensitive subject for any future machine learning and artificial intelligence development due to the capacity of AI to capture minute details and behavioral patterns that may not be obvious to the users themselves.

ETHICAL RISKS IN THE CROSS SECTION OF XR, GIS AND AI

If artificial intelligence (AI) is the next evolution of our economies and the driving force behind the future constructs of the world, then it stands to reason that from its onset practitioners should strive to make it fair, equal, open, and non-judgmental. To include as much context and kindness into every algorithm, to provide the future generations with the toolbox for a more opportune future.

As extended reality (XR) technology is the next evolution of human interactions with information, it stands to reason practitioners have an ethical responsibility to make it transparent, understandable, and safe for users of all ages, geographies, and social groups (Forbes Technology Council, 2021).

If GIS is to remain an evolving scientific field of innovation and advanced digitalization it ought to cohesively capture not only the economic opportunities of modelling, but the ethical risks and flows within its comprehensive set of model frameworks.

Ethics in the realm of science today is much more complex than it has ever been, as data science evolves to capture more of humanity in its clean-cut definitions, which often translate to real life implications. Ethical risks in the convergence of digital fields will continue to pose challenges for practitioners, who are striving to achieve quality of service and efficiency. Since cross sections are more difficult to regulate due to the variety of frameworks they could fall under and the variety of loopholes in the otherwise growing legal AI realm, these will become an even more expansive field for analysis and scenario examination.

Ethical risk in the cross section of XR, GIS and AI could be examined in the two broad categories, which resonate with the broad categories for analytics in artificial intelligence but contain an expansive plethora of other factors.

Risks in Environmental Interactions

Environmental interactions are the ones where the geographic information systems (GIS) aspect of the ethical risks is most pronounced.

Immersive technologies with the geo element in them represent a risk for all actors involved in harmful scenarios as datasets with GIS components comply by the basic topology rules of each object being unique, identifiable, and connected to other unique objects in specific ways. The utilization of that aspect of the technologies would make identification much easier than it is in the current digital economy, where means still exist to remain anonymous. Hence, caution should be applied for the protection of the physical well-being of users as well as their emotional well-being. But the more comprehensive risks lie in the obvious ethical dilemmas for the various players in the marketplace – on how to best apply the technology in a conscientious manner, which includes within itself the matter of ethics not as a standalone concept outside of the product, but as an integral part of it.

There are requirements and guidelines on how to develop and design sustainable products, there are disclaimers on data privacy in effect upon purchase of the most basic ocular set. There are even guidelines on how to best utilize the capabilities of computer vision to produce balanced and adaptive results. The general criteria and the regulatory aspects could be comprehensively covered by company policies, but there would still be risk scenarios for examination in every instance of the AI application with the product.

What needs to remain in sharp focus is that AI is still imperfect and reliant on imperfect data and when applied with raw inputs from ocular devices, it is bound to produce majority dependent results – to play on the passions and emotional stimuli of the persona, to exaggerate routines and behaviors, to predict a preferred sight seconds before you look at it.

The scenario of experiencing an augmented reality (AR) enhanced tourist attraction, a virtual walk, or an entire day where marketing occurs entirely within glasses is already here (Shah, 2019) and the very design of that experience could have lasting impacts on industries, locations, and people.

Object-oriented modelling is a part of various aspects of data science. Geographic information systems (GIS) model events in the world to hypothesize about potential scenarios, develop solutions and contextualize events in interlinked relationships, within which no human being is an island on its own.

Predictive analytics would simplify daily tasks, but it would require the gathering of geolocation data and tracking the behaviors of users not only within the boundaries of their online environment, but within the real world as well. The above is sensitive information for a variety of reasons – social and economic profiling, gender profiling, profiling based on geolocation.

Modelling algorithms classify users within certain demographic groups, professional, social, and shared interest environment (Martin, 2019), which enhance the spend and efficiency of available solutions, but depending on the social interaction context and virtual context these solutions often result in bias bubbles and "birds-of-a-feather" modelling. Therefore, it is important for the future XRGIS designers to be aware of the bias flaws that the driving algorithms of their solutions might carry and to strive to provide as balanced an offering of services as available within that context.

A city layout is but an expansive framework of relationships and topologies within which every daily habit and pattern could be monetized for the benefit of vendors and advertisers within the paradigm of a new extended reality economy. Urban realities would easily become a playground for the subliminal tourist experience and the capture of additional data points.

The sensemaking aspect of XR and AI cross sections could accelerate the effect on the emotions of entire groups of individuals in relation to tourist attraction (ambiance and associations), smart city "digizen"[4] dynamics and the interactions between individual actors and senses within real life environments (geodata in relation to IoT sensors, smart buildings, and urban design elements).

The ways in which citizens perceive their built environment will be captured and analyzed for the overall improvement of urban ecosystems and the service industry. Curated content would be easily projected before the user with interactive billboards and urban designs element, based on preferences and tastes gathered by ocular technology. If based on purchase power classification, cultural and emotional preference, the experiences of a city for its variety of social groups and classes could be vastly different.

These elements would be easily deployed in the GIS model pipeline to produce predictions, outcomes and scenarios weighted in one direction or another. The dilemmas in that real world application of these technology conversions would produce a set of brand-new set of risks for the various players in the digital marketplaces.

For city officials it will be what types of data is being captures in the urban ecosystems to ensure the physical safety of the citizens, for the marketers it will be a brand new competitive space for attention capture and a new market for sensemaking advertising, for the users it could become a question of which reality is the real one as even now we are experiencing the effects of deep fakes in digital world even without the enhancement of virtual realities with the added senses of vision and hearing in 3D.

Humans interpret the world through multiple lenses, determined by both culture and environment, which form the basis for human bias and deeply influence how people interpret their relationships and sense of self. The nature of these lenses and how reality is experienced has dramatically evolved with the introduction of digital technologies and easy access to information via the web. Our perspectives are profoundly influenced not only by the cultural values of a global population, but by the underlying tracking technologies fueling the economic underpinnings of the web." (Institute of Electrical and Electronics Engineers [IEEE], 2021, para. 2)

Marketing algorithms in the current structures of the internet are built on AI-based conceptualization of the human being in front of the blue screen. Subliminal suggestions in extended reality would easily curate to a sensemaking craving for products. These are powerful tools for growing a brand-new design economy that captures every preference and utilized experiences rather than advertisements. But a 6.0 virtual economy (Christensen et al., 2016) which is regulated by AI-modelled definitions could produce adverse effects on individual experiences and could even be weaponized against various personality types, ethnic and social groups, producing vastly different experiences for the groups of “digizens” in the smart city.

There are two perspectives, which need to be covered in further studies about environmental cross section risks.

One is the service vendor perspective, which could be additionally explored in a digital twin scenario, where users are presented with different options of the city design elements and commercial offerings of various services, which could transform the standing of a business or the experience of a tourist site. The other one is the user perspective on how they want to be treated within that new market environment. There are applications in the current marketplace, which utilize augmented reality to offer consumer services within the makeup and fashion industries, there are simulation applications of a second world experience and expansive opportunities for further engagement within the consumer markets of the food and beverage industries, where the experience of thirst and hunger would be much easier to convey. Those would significantly transform the world of commerce, but would likely lead to additional gathering of data from customers, which in the instance of gaze and iris tracking might even be subconscious.

An advanced sensemaking 6.0 digital economy is likely to transform daily interactions in unimaginable ways as it is set to introduce a brand-new set of technology applications, enhanced with AI integrations. What this would mean for the routines and actions of human beings in their extended reality is another element of the ethical risks of the proposed dilemmas as both vendor and end-user scenarios in the new economy will be hard to predict and regulate.

Risks in User Interactions

AI algorithms are capable of capturing various behavioral tricks and patterns, of transforming phrasing from a NLP(natural language processing) feed into a psychological profile and providing an analytics of the type of personality beyond the social media profile (level of introversion or extroversion, or neuroticism). These algorithms are utilized for the purposes of digital personality hacking.

Various sales techniques are based on personality hacking from basic person to person barter sales to social network engineering and online marketing. Personality hacking is possible when there are enough data points about an individual to develop a picture of them, of their behavioral and commercial patterns, of their characteristics and interactions.

These data points could consist of life events, impressions, preferences and set of behaviors under a variety of scenarios. Adding these data points to a vendor data stream might transform the experience of a user far beyond the digital and into the real world.

Social networks and search engines utilize a form of personality hacking based on preference scores and metrics. Advertisements and search content curation are based on user modelling criteria, which produce vastly contrasting user experiences in the online world. The results of a search engine in one region of the world would be vastly different from those in another, depending on the most popular news articles of the country and the most popular music in the region. Hence, the ways in which the world is viewed vary region to region, nation to nation, person to person and internet services are running against the task of its creation to unite by solidifying differences based on aggregation and suggestion metrics.

Extended reality technologies (XR) which includes ocular devices and augmented reality content maintains the likelihood of transforming the perceptions of various user groups in the real world.

The addictive nature of digital realities could represent a further ethical dilemma for the regulations that need to be attached to purely commercial sensemaking scenarios as they would rely mostly on derived, even subliminal information (gaze capture), difficult to translate within the scope of a singular legislative framework.

Internet addiction is a chronic, global problem with users from all age groups and regional geographies experiencing the world in much more digital manner than ever before. The most addictive forms of online escapism, including YouTube, TikTok and Netflix utilize AI in their backend to provide seamless experiences to their users. These applications are with continuously growing audiences that spent hours of their lives submerged in that form of entertainment. An XR element added to the real world has already exposed the addictive nature of these applications with the PokemonGo trend, where an overlay of animated characters was attached in interplay with locations in the physical world.

The connection between the augment element and the real world might be too powerful to overcome if the curated content improves in quality to the level in which the differences between what is real and what is augmented is hard to recognize as utopias would be hard to come by in the context of the reality but would be easy to curate in the extended reality version.

Extended reality headsets are being utilized in several studies as a solution to alter the psychological experiences of users in fundamental ways – for the treatment of addictions and phobias (Slater et al., 2020). If XR technologies could be utilized to regulate emotional responses, these would immediately become a significant risk vector when enhanced with AI-curated content and geo interactions. Both the AI-curated content models and geo interaction models have already been explored by online services in the purely commercial sense. But the fabric of ethical dilemmas would thicken when additional senses are included in the experience as the brain can often be tricked by sight. To complicate the matter further, vendors are already considering extension of the sense of sight in XR to include gloves that can simulate touch and sound overlays in the already existing headsets. This points to an extensive move towards a new way of communication, which will be conducted via online means, but will remain a matter of the senses.

There will be so many data points gathered in such close interaction between the human being and the hardware, that XR would become the natural breeding ground of next generation AIs performing quicker computations with more data points.

The capacity of these technologies to affect a person's experience and emotions would have a lasting effect on both user-user interactions and the user-environment interactions.

The "avatarization" of human beings (Takahasi, 2020), where the perception of a being is altered to a more digitized and surreal version of living being, poses further risks as demonstrated by exponential growth of online hate culture, except with XR, actions would have even more pronounced harmful ramifications due to the sense engagement of the devices. While XR technologies in cohesion with AI deep fakes are being utilized for educational exercises to decrease implicit bias behaviors, a recent study by the University of Barcelona (Banakou et al., 2020) has discovered that the "avatarization" of human beings in virtual reality scenes could lead to quite the opposite – contextual development having adverse effects on racial perception.

Within the context of an AI-content curated extended reality, which is further connect to experience of social groups within a city or within any other real-world context, there are risks for exacerbation of the negative aspects of social network engineering currently being experienced in the digital world in particular group think, silos and hate chambers. If human beings only exist within the bounds of curated content, generated around their preferences, then their biases would solidify (Dickson, 2019).

Bias bubbles within the cross section of XR, AI and GIS would have very real implications and would be very complicated to rectify as these technologies engage sensemaking experiences, which affect brain and memory, and cognition.

The ambiance, feeling and experience of a locale remain with tourists for years after a visit. The feeling of hate and fear within a virtual reality scene would have the capacity to affect users in fundamental ways, which is likely why these technologies are successfully utilized in the medical domain to combat phobias and to train emergency personnel. While these applications are revolutionary on their own they run the implicit risk of enhancing negative emotions or the negative side effects of addicting positive experiences.

AI technologies have the capacity to identify emotions from micro expression, from mimics and tone of voice. There are extensive studies being conducted in emotional reflectance in robotics to enhance acceptance of the suggested solutions. XRs in combination with an AI backend, which is capturing the subconscious preferences of its users might create a powerful preferable set of realities for the engaged users. As a tool such a solution would be a revolution for marketing, for economics and for a new service industry where the physical and the digital worlds are becoming more and more interchangeable.

Emotional vulnerability is a human condition not particular to a specific ethnic, social, religious or gender group. Psychology and sociology recognize that there are more vulnerable age groups, conditions, and environments, within which human beings are not completely able to rationalize the emotions they experience, where the perceptions and views acquired could leave marks, triggers, alter or completely transform the trajectory of a person's life.

The introduction of a deep fake element in the context of a scene could result in harmful side effects for various groups – especially for mentally vulnerable individuals, the elderly, and children.

A loved one brought back to life within an artificial environment or a deep faked human being, capable of emotional influence within these environments is not beyond the realm of possibility. The deep fake avatar of a lost loved one could be extremely difficult to overcome as it relies on universal human vulnerability and should be treated as an extreme risk from all vendors aiming to provide AI curated XR services. While the existence of well-designed avatars has the capacity to alleviate loneliness and emotional pain in certain scenarios, there could be significant ramifications to the fabric of society –

preferred artificial connection than a human-human, preferred extended reality to the real world, deeper emotional dependency on technology.

CONCLUSION

'These are the best of times, these are the worst of times' for immersive technologies and AI, as those are far more attractive than they've been with the receding Covid-19 pandemic, while the responsibility for their adoption within numerous use cases has never been greater, considering the implications of a much more divisive internet space.

Every single innovation has its testing phase, its downturn, its issues and bugs, its problems to maintain and its unintended consequences that often affect the most vulnerable user groups (Todt et al., 2018).

The scenarios examined above are based on existing studies and current applications of these converging technologies. The aforementioned cases represent the broadest classification available within the AI and XR fields. Further cases will arise as human experiences continue to evolve along with the emergence of far more diverse XR applications.

While XR and GIS, and AI appear to not have that much in common as they are the paradigms of different sciences and focus of different players in the digital space, and even the amalgamation of different coding libraries, in the current digital space it is the intersections, which produce the most significant innovations and interesting opportunities to explore and analyze, and 'play' with the opportunities. But *'with great power comes great responsibility'* that requires the further analysis of every single fail scenario, where these technologies are applied.

The accelerated digitalization of the current data generation is conducive to the rise of extended reality technologies, enhanced machine learning, social network engineering and the development of new economic constructs, which could have unintended side effects. While these innovations on their own are impressive, their combined effects would once more transform the daily existence of the regular human being.

Data science and innovation practitioners within the field of AI, need to maintain integrity in the face of quick and easy business solutions and to treat every user of the emerging technologies as a multi-dimensional human being and not a number on a dashboard, removed from implications outside of the digital domain.

Integrations of AI within appliances will supply convenience and improve experiences, and XR is set to become one of those appliances that are an inextricable part of daily routines. Scenarios of potential application span from enhanced communications with 3D graphics and 'teleportation', from experience curation via artificial overlays in a geolocation, to engagement of the senses in new ways by transforming an environment - that is a technological revolution in the purest sense.

Regulations and standards will undoubtedly grow to reflect the needs for privacy, security, and equality within the cross sections of these technologies, but practitioners are obliged to consider and contextualize the broader ramifications for every set of recommendations, sensemaking and curation of content within the individual experience.

Environmental and user centric factors should be considered from all points of view, a variety of disciplines and professions, capable of recognizing the validity of emotion and individual experiences in the context of extended reality in the AI-powered global digital economy.

The suppliers of the AI should reflect on the side effects of the solutions in the concrete scenario as with XR solutions there are use cases and data streams vastly different from the ones in autonomous vehicles. The coders need to understand the ethical blind spots of the efficiency algorithms that result in their product outputs. The users need to be aware of the invisible data streams that are likely to be captured by a more sophisticated AI algorithm.

Whilst the exponential growth of reference data points suggests a higher level of sophistication in their processing, the outputs of the algorithms are too simplistic to capture the wider scope of opportunities within the physical world, which leads to a need for a greater amount of care in applications where the line of distinction between the realities is getting blurred. While XR designers are mostly occupied with the speed and design of the experience, there needs to be a consideration for the backend enhancements that go along with the beautiful and ambient frontend. Artificial intelligence carries the perspectives of the future as a transformative wave-effect set of innovations which will transform almost every single economic segment, user group and industry it touches. The level of reliance on algorithmic outputs will become an instinctual part of reality as the suggested data inputs intermingle with preferences and external biases.

In a vastly more digital world, extended realities will supply enhanced experiences that touch on the senses, promote deeper connections, and provide value to end users' digital services like they have never experienced before. All of it leading to a new economy of market and experience, that will fundamentally affect its user groups – from new ways to learn, to new ways to shop, to new ways to communicate – all of which in cohesion with AI.

As further use cases arise, the need for ethical approaches in the design of both environmental and user-interactive application of XR will grow and transform, and there are additional cross sections, which need to be examined considering the constantly evolving designs of technologies and their plethora of applications – there are potential intersections with sociology and economics, and marketing, and psychology.

The future of geographic information science has never been more interesting, because it is an evolving science, where terms like 'topology' and 'location' are garnering new meanings in the variety of cross-cutting sciences that geography interacts with. There is a micro and a macro aspect to GIS becoming more significant with every examination of the science. If an individual is considered within the realm of data science as a set of data, within GIS an individual is a subject in an unique topology of interactions and scenes inextricable from their physical world implications. The topology of XR headset could relate both to the private microcosm of the individual and to the wider network of users and user integrations on a global scale.

Further interesting cases for analysis will be the scale and scope of these changes as they begin to occur in the daily experiences of digital service users in the next twenty to fifty years. The fundamentals of these environment and user centric interactions are being laid out right now, but the changes will be a field for exploration in years to come.

On the macro scale, the issue of a smart city digizen application of XR/AI/GIS intersections will be profoundly important for the future of advertisers and service providers in the interconnect city scape.

On the micro scale, the ramifications of AI enhanced user interactions in an augmented high-quality reality will supply a variety of interesting use case scenarios.

Within these changes the future generations will also change and even the perception of what ethics is and ought to be might transform to reflect the premises of this 6.0 economics of the senses.

FUTURE STEPS

The ethical dilemmas attached to the application of within the cross section of extended reality, geographic information systems and artificial intelligence need further exploration. Extensive empirical studies should be developed around the various applications of XR sensemaking in the context of a smart city. While easy to ignore in the periphery of all other fundamental risks of XR applications, the experiences of various demographic groups in the city need to be considered as AI curated content already draws lines of separation in the digital world that should not be affirmed in the real world aspect.

AI is still in its beginning stages of development, where the data inputs are easy to track and the logic of a solution as well as its faults is easy to follow. As algorithms working in the background become far more sophisticated and data inputs grow exponentially, it will become much more challenging to follow the ‘plot’ and context of suggestions.

While the premise of a sensemaking XR-based 6.0 economy seems far off in the distance, the 2020 of accelerated digitalization has only proved that humanity is becoming more and more reliant on its technologies to supplement physical world experiences. These processes have been present in the generational ratio of changes in the digital domain, but as accelerated by external forces they are becoming even more apparent.

If social networks further their growth into the fields of extended realities, we could rapidly get accustomed to the experience of connecting in 3D environments, of an enhanced geo world where the topology of the digital and the physical are no longer separate.

The ethical dilemmas within these converging instances of digital solutions will require an additional set of standards and approaches as new applications arise. An integral bit of work remains to be developed for the GIS aspect, as it is relevant for both the physical and the emotional implications of the XR economy. The constant risk of bias amplification within the realm of AI needs to also be addressed and acknowledged by the variety of market players involved in the development of XR solutions.

The future of XR is exciting for both its developers and its users and ethics and standards will take nothing away from that but will supply additional layers of security in the growth of the new economy. Users should be able to enjoy the benefits of a sense market without worrying about the ways in which their personalities are being exploited by an underlying AI. Digizens should be able to enjoy the pleasures of digital city art without the worry of embedded imagery which could have a subconscious effect on their day.

There are many more positive aspects than there are risks in the applications of XR/GIS/AI. But for every player in the marketplace to be able to take full advantage of the positives, there is also the need for ethical applications. The capacity to see and comprehend the unintended side effect, the capacity to recognize and counteract biases in all their forms, the capacity to develop robust algorithms that do not rely on simplistic classifications – these will be fundamental for the successful growth of XR.

It is also fundamental to remember that ethics is much easier to apply in a single domain, where there may be previous experience and use cases, and a playbook, and a set of rules to follow in terms of what should be done. The challenges for ethics will expand as the areas of digital convergence grow and the scenarios diverge to reflect individual use scenarios.

ACKNOWLEDGMENT

Most sincere gratitude to Richard Foster-Fletcher, chair of MKAI for including me in the working group on XRSI Safety Standards, which provided the necessary background for this article.

REFERENCES

AI-Driven platform for content curation — or a Netflix of knowledge! (2017, July 16). *Bold Business*. https://www.boldbusiness.com/human-achievement/ai-driven-platform-content-curation/

Arisona, S., Mueller, P., Meriaux, A., & Hansen, R. (2020, March 10-13). *Extended reality (XR) with ArcGIS* [Paper presentation]. 2020 Esri Developer Summit, Palm Springs, CA, United States. https://www.esri.com/content/dam/esrisites/en-us/events/conferences/2020/developer-summit/extended-reality-xr-with-arcgis.pdf

Artificial intelligence. (2021, August 3). In *Wikipedia*. https://en.wikipedia.org/wiki/Artificial_intelligence

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., & Slater, M. (2020). Virtual body ownership and its consequences for implicit racial bias are dependent on social context. *Royal Society Open Science*, *7*(12), 201848. doi:10.1098/rsos.201848 PMID:33489296

Bardi, J. (2018, November 11). *3 secrets to creating immersive virtual environments with Unity and Vuforia*. Marxent Labs. https://www.marxentlabs.com/virtual-environments-unity/

Berry, W. (1971). *The unforeseen wilderness: An essay on Kentucky's Red River gorge*. Counterpoint.

Christensen, L. R., Marcik, W., Rafert, G., & Wong, C. (2016). *The global economic impacts associated with virtual and augmented reality*. Analysis Group. https://www.analysisgroup.com/globalassets/content/insights/publishing/analysis_group_vr_economic_impact_report.pdf

Çöltekin, A., Griffin, A. L., Slingsby, A., Robinson, A. C., Christophe, S., Rautenbach, V., Chen, M., Pettit, C., & Klippel, A. (2020). Geospatial information visualization and extended reality displays. In H. Guo, M. F. Goodchild, & A. Annoni (Eds.), *Manual of digital earth* (pp. 229–277). Springer. doi:10.1007/978-981-32-9915-3_7

Committee for the Coordination of Statistical Activities. (2020). *How COVID-19 is changing the world: A statistical perspective* (Vol. 1). https://data.unicef.org/wp-content/uploads/2020/07/covid19-report-ccsa.pdf

Dickson, B. (2019, May 20). *Artificial intelligence created filter bubbles. Now it's helping to fight it*. TechTalks. https://bdtechtalks.com/2019/05/20/artificial-intelligence-filter-bubbles-news-bias/

European Commission. (n.d.). *Data protection: Rules for the protection of personal data inside and outside the EU*. https://ec.europa.eu/info/law/law-topic/data-protection_en

Feldman, J. (2020, October 18). *Flickplay's 3d social media platform presents as an industry first*. Influencive. https://www.influencive.com/flickplays-3d-social-media-platform-presents-as-an-industry-first/

Forbes Technology Council. (2021, January 6). Tech experts predict 13 areas AI and VR are set to revolutionize. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2021/01/06/tech-experts-predict-13-areas-ai-and-vr-are-set-to-revolutionize/?sh=3cd7283b25b0

Institute of Electrical and Electronics Engineers. (2021). *The IEEE global initiative on ethics of extended reality*. IEEE Standards Association. https://standards.ieee.org/industry-connections/ethics-extended-reality.html

Johns Hopkins University. (2021). *COVID-19 dashboard*. Retrieved August 3, 2021. https://coronavirus.jhu.edu/map.html

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. doi:10.100710551-018-3921-3

Overby, S. (2020, May 7). *5 artificial intelligence (AI) types, defined*. The Enterprisers Project. https://enterprisersproject.com/article/2020/5/5-artificial-intelligence-ai-types-defined

Publications Office of the European Union. (2020). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence act) and amending certain union legislative acts* (COM/2021/206 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

Scribani, J. (2019, January 16). *What is Extended Reality (XR)?* Visual Capitalist. https://www.visualcapitalist.com/extended-reality-xr/

Shah, M. (2019, August 21). *How Augmented Reality (AR) is changing the travel & tourism industry*. Towards Data Science. https://towardsdatascience.com/how-augmented-reality-ar-is-changing-the-travel-tourism-industry-239931f3120c

Sirk, C. (2020, August 21). *Diffusion of innovation: How adoption of new tech spreads*. CRM. https://crm.org/articles/diffusion-of-innovations

Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., Watson, Z., Breen, G., Schwarz, R., Steptoe, W., Szostak, D., Halan, S., Fox, D., & Silver, J. (2020). The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality*, *1*, 1. Advance online publication. doi:10.3389/frvir.2020.00001

Takahashi, D. (2020, October 27). *Alethea AI makes it easy to create AI avatars from a single photo*. VentureBeat. https://venturebeat.com/2020/10/27/alethea-ai-makes-it-easy-to-create-ai-avatars-from-a-single-photo/

Todt, G., Weiss, M., & Hoegl, M. (2018). Mitigating negative side effects of innovation project terminations: The role of resilience and social support. *Journal of Product Innovation Management*, *35*(4), 518–542. doi:10.1111/jpim.12426

United Nations. (n.d.). *Millennium development goals*. https://www.un.org/millenniumgoals/

KEY TERMS AND DEFINITIONS

Artificial Intelligence (AI): Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

Data Science: Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.

Digital Economy: The digital economy is the worldwide network of economic activities, commercial transactions and professional interactions that are enabled by information and communications technologies (ICT).

Extended Reality (XR): Extended reality (XR) is a term referring to all real-and-virtual combined environments and human-machine interactions generated by computer technology and wearables, where the 'X' represents a variable for any current or future spatial computing technologies. It includes representative forms such as augmented reality (AR), mixed reality (MR) and virtual reality (VR) and the areas interpolated among them. The levels of virtuality range from partially sensory inputs to immersive virtuality, also called VR.

Geographic Information System (GIS): A geographic information system (GIS) is a system that creates, manages, analyzes, and maps all types of data. GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). This provides a foundation for mapping and analysis that is used in science and almost every industry. GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication and efficiency as well as better management and decision making.

Sensemaking: Sensemaking or sense-making is the process by which people give meaning to their collective experiences. It has been defined as "the ongoing retrospective development of plausible images that rationalize what people are doing" (Weick, Sutcliffe, & Obstfeld, 2005, p. 409). The concept was introduced to organizational studies by Karl E. Weick in the 1970s and has affected both theory and practice.

Virtual Reality (VR): Is a simulated experience that can be similar to or completely different from the real world. Applications of virtual reality include entertainment (e.g. video games), education (e.g. medical or military training) and business (e.g. virtual meetings). Other distinct types of VR-style technology include augmented reality and mixed reality, sometimes referred to as extended reality or XR.

ENDNOTES

[1] Referenced artificial intelligence classifications The artificial intelligence (AI) classification, based on capacity is dependent on computational capacity: 1. Narrow – capable of solving specific tasks 2. General – capacity to solve as many tasks as a human being 3. Super – surpassing the capacity of a human being to solve tasks The artificial intelligence (AI) classification, based on functionality is a far more interesting one, which encompasses: 1. Reactive – solving problems upon prompts 2. Limited memory – running the same algorithm to solve similar issues 3. Theory of mind – development of EQ (emotional intelligence) algorithms 4. Self-awareness – the AI knowing it is an AI.

[2] With the cardboard versions of extended reality headsets available in the EU and UK at a price point of 1,5 EUR per set.

[3] The usage of XRGIS within the article is intentional to highlight that positioning and GIS is not only applicable in the VR instance, but in all other instances of the technology.

[4] Digizen in the article refers to a citizen involved in the digital economy in their area, involved in the exchanges and solutions being provided at scale as both a producer and a consumer of a data service.

Chapter 15
Discussions on How to Best Prepare Students on the Ethics of Human–Machine Interactions at Work

Cynthia Maria Montaudon-Tomas
https://orcid.org/0000-0002-2595-6960
UPAEP Universidad, Mexico

Ingrid N. Pinto-López
https://orcid.org/0000-0002-1580-1375
UPAEP Universidad, Mexico

Anna Amsler
https://orcid.org/0000-0003-3183-0878
Independent Researcher, Mexico

ABSTRACT

This chapter analyzes the evolution of the new ways of working, especially in terms of algorithms and machine learning. Special attention is given to algorithmic management and its ethical concerns, as well as to practical examples of the application of algorithms in different sectors. Faculty discussions about how to best prepare students to deal with human-machine interactions at work are presented, with algorithmic management and accountability the discussion's central axis. In algorithmic management, there are distinct positions to analyze; one that favors innovation and efficiency and privileges dignified work and ethics. A brief proposal on introducing algorithmic ethics into the programs offered at a private business school in Mexico is included.

DOI: 10.4018/978-1-7998-8467-5.ch015

INTRODUCTION

Advances in technology, especially in Artificial Intelligence (AI), have been exponential. Digital innovation touches every aspect of life, both personal and social, affecting how we understand the world and ourselves. Digital innovation and AI are increasingly present in human activity and decision-making and are altering the way we think and act. A study by PwC has even suggested that human-machine interactions will soon become as common as watercooler conversations between colleagues today (Donkor et al., 2017).

AI is driving the adoption of technology at an unprecedented rate (Oracle, 2019). The application of AI in the workplace is being led by India, China, the United Arab Emirates, Brazil, Australia, New Zealand, Singapore, the United States, the United Kingdom, France, and Japan (Oracle, 2019). Organizations are increasingly using AI to manage their workforces through algorithmic technologies that rely on large datasets, making it possible for machines to become bosses (van Hooijdonk, 2019).

In this regard, employers are ultimately responsible for their employees' wellbeing, and the underlying issue is whether or not algorithms will be able to treat people fairly. AI and algorithms offer many opportunities to design more flexible, fulfilling ways to work, but they need to be developed and managed ethically and effectively (Walsh, 2019), and it is crucial to understand how this technology makes decisions (Heilweil, 2020).

Numerous questions require a prompt response, as human-machine interactions will continue to increase, especially those regarding employees' ethical responsibility following orders and decisions produced by an algorithm or a robot. Others include how socialization at work is going to be affected, the essential changes we will see in terms of requirements of educational programs, and more specifically, how to ensure that students will consider AI as a support to human decision-making and action, instead of promoting a lack of accountability and the loss of free will.

There are no easy answers, and machine learning in itself is not unethical. Human identity and dignity will be impacted based on how algorithms are developed, and it is the interaction between machines and humans that needs to be addressed.

The field study is centered on a private higher education institution in central Mexico. General conditions regarding stress and working hours in the country are described to create the general background of the study, along with two significant regulations that legislate psychosocial risks and remote work. The population considered were full-time faculty members who had moved their activities online. A section of a scale that is part of a more extensive study about the effects of remote work during the pandemic was used. Results were analyzed as a whole and later on divided according to gender to determine whether there were significant differences in terms of burnout syndrome in faculty.

BACKGROUND

The New Ways of Working, Algorithms and Machine Learning

The phenomenon of *datafication* (Jarrahi & Sunderland, 2019), based on increased technological change and the emergence of NWoW, has created unprecedented scenarios from which to analyze labor as the center of the problems when examining technology in the ecosystem of human work. Displacement of work by machines is a critical aspect of analysis now that thinking machines are replacing human labor.

The exponential rate of new developments, particularly AI and its algorithms, has left human decision-making out of the equation. Task or operations are now being done through a particular form of *intelligence*, something humans would otherwise carry out.

Until recently, work was considered an intrinsically human action, and only people were capable of work. Nowadays, technology is expected to become an ally at work and an instrument to make it more effective. However, machines are surpassing human analytical capacity; work can become dependent on technology, and humans are missing the opportunity to act as persons at work (Alford & Naughton, 2001). Algorithms are impersonal by nature; they are based on large amounts of data, specific programs, and general assumptions, which leads to the depersonalization and dehumanization of what once was human work, sometimes absolving humans of their responsibility.

Algorithms are step-by-step coded procedures that determine how tasks on the internet are performed (Bell, 2001); they are used to solve a problem and are usually expressed in computer codes as a set of instructions to follow to complete a task (Kritikos, 2018). The computational formula autonomously makes decisions based on statistical models or decision roles without direct human interventions (Fernández-Macías et al., 2018). Algorithms are presented as objectively and mathematically correct (Jago, 2019), and they rewrite themselves as they work. Self-learning algorithms are now making and executing decisions affecting labor, limiting human involvement and oversight of the labor process (Duggan et al., 2019).

However, there is a lack of transparency in using algorithms since AI involves machine learning, and it develops its models to make assessments and decisions, making it virtually impossible to identify the path taken. Machine learning finds reproducible patterns in data; these patterns have predictive power in that they can anticipate the target value for new and unknown inputs, but rather than being universal and objective, it produces knowledge that is irrevocably entangled with specific computational mechanisms and the data used for training (McQuillan, 2018). They are complex techniques where it is hard to decipher precisely how each input drives model outcomes, often resulting in them being characterized as “black boxes” (Bigham et al., 2018; Loggins, 2020).

Algorithms have been considered the intellectual capital of the internet (Bell, 2001). Among their advantages is reducing time spent on repetitive and time-consuming tasks, allowing people to focus on human empathy and judgment, which matters the most (Batterywala & Agarwal, 2018). The goal of automation is to increase efficiency (Roose, 2019), and as new tools have been created, the labor market has adapted to include these developments to further productivity.

Initially, it was thought that management roles were relatively safe from automation. However, the use of algorithms, AI, and robots has broadened. In trading and banking, human participation is diminishing, being replaced by computers, algorithms, and passive managers (The Economist, 2019). As an example, algorithms have been around for decades in portfolio management. Algorithms determine the work schedules in retail hospitality, food and services, and other industries. Long and unpredictable hours are increasingly being scheduled by algorithms that analyze seasonal sales patterns, customer trends, and even the weather (Loggins, 2020). They are also watching over hotel housekeepers, telling them which room to clean and tracking how quickly they do it. They listen to call center workers, telling them what to say and keeping them busy (Dzieza, 2020).

Hiring platforms, such as LinkedIn, use algorithms to sort through thousands of profiles to recommend job candidates to company recruiters (Carey & Smith, 2016). In this sense, algorithms enable digital labor platforms to automatically manage transactions between thousands of gig workers and service recipients (Jarrahi & Sutherland, 2019).

Companies are using critical intelligence and other elements of the work-on-demand business model to automate workforce management tasks. Algorithms are increasingly making managerial decisions that people used to make (Lee, 2018), rewriting their work roles (Rosenblat, 2018). This context has raised a question regarding whether software can be better at managing people than human managers.

Algorithms are at work on just about everything, including analyzing genes for diseases, life insurance, and more. Algorithms are silently structuring our lives, determining whether someone is hired, promoted, offered a job, or even provided housing (BBC, 2021; Martin, 2019; Polzer, 2018). Algorithms mediate social processes, business transactions, and governmental decisions; they influence how we interact among ourselves and with the environment (Mittelstadt et al., 2016). They can drive everything from browser searches to medical care. In the workplace, they assist in various tasks, including measuring productivity, evaluating performance, and even terminating employment (Barratt et al., 2020).

Machine learning has been rapidly adopted for research and discovery across academia, business and government, and it is becoming a kind of dark matter that invisibly distorts the distribution of benefits and harm (McQuillan, 2018). Advances in AI, machine learning, and data infrastructure are transforming how people govern and manage organizations; algorithms are being used by policymakers, physicians, teachers, police, labor platforms, and others (Lee, 2018), expanding their reach into and throughout human life.

Critics of companies using algorithms for management tasks suggest that automated systems dehumanize, unfairly punish employees, and provide an environment with no privacy because they can track everything employees do. Problems like these have become a source of tension between workers and the platforms that connect them. Furthermore, the use of algorithms in supervising positions can have other downsides: they can be biased from their design, the decision-making process is opaque, and its use sometimes results in a feeling of isolation.

In a new world of work where subjective judgment is compiled and quantified by technologies, human activities become subject to metrification, classification, comparison, and market competition (Curchod et al., 2020). Humankind is transforming into an instrument of progress. Human actions are becoming determined by machines, consequently taking away dignity for human work. Human action cannot depend exclusively on what a machine has determined, undermining traditional figures of authority.

Machine thinking is creating decisions without humans knowing how the decisions are being made. Machines are extracting information based on logical statements to encounter more complex situations, but the underlying principles for those decisions are not written into computer codes; therefore, humans have no possibility of understanding whether a decision or rule is ethically correct or not.

While most scientific debates center on the progress required to establish what is relevant for AI to reason ethically, research is needed to settle how we should be teaching students to work in an ever-increasing technological environment in which AI will most certainly have leading and arbitrage functions to which graduates will be subjected in their everyday actions and decisions while on the job. Business schools' courses should include materials on the consequences of AI in human resources, business models, markets, social security, finances, health, and others, as well as the ethical ramifications of choices that can appear to be ethically neutral (Burton et al., 2017).

Algorithmic Management

Algorithmic Management (AM) is a term coined by academics at Carnegie Mellon University from the Human-Computer Interaction Institute (O'Connor, 2016). It has been presented as one of the recent dilemmas for leaders in the age of AI, determining when and how to use algorithms to manage people and teams (Walsh, 2019). It includes behavioral economics and further efforts to manage or manipulate human work (Capelli, 2018), utilizing technology's influential role to monitor, manage, and control workers (Rosenblat & Stark, 2016).

Management by algorithms is a diverse set of tools and techniques to manage workforces, relying on data collection and surveillance of workers to enable automated or semi-automated decision-making (Mateescu & Nguyen, 2019). In this sense, algorithms and not humans decide how business operations should be performed (Walsh, 2019).

Machine learning emerged alongside the rise of the gig economy (Capelli, 2018), and it is participating more and more in activities previously considered uniquely human (Silver et al., 2016). The gig economy is a growing component of the global labor market. It is disruptive to the traditional understanding of work because it is based on a form of contingent labor such as the on-demand or work-as-required. Across all gig works, the common denominator is the presence of an intermediary in the form of a digital platform organization (Harris, 2017). Algorithms manage, track, discipline, and set expectations for workers without human supervision (Duggan et al., 2019).

AM is essential to platforms. It has a significant bearing on the overall experience of gig workers (Jarrahi & Sutherland, 2019). Platforms in fields such as ride-sharing and food delivery are built on algorithms (Captain, 2020). What is going on is mainly invisible to workers because the platform obscures all the details, and the invisibility factor can lead to abuse (Moor, 1985) and can also lead to data tampering and even unrecognizable bias.

AM is based on four broad areas: technology-enabled surveillance and control, transparency resulting in power imbalances, bias and discriminatory practices in the workplace, and accountability, due to the obscuring of specific decision-making processes (Mateescu & Nguyen, 2019). It presents numerous benefits, for instance, a lack of favoritism, no personality clashes at work, no yelling, no bad moods or harassment, no ambiguity, only prescriptive instructions, all work done on schedule, and, a plus during pandemic times, algorithms will not get sick or need to wait for a vaccine.

Furthermore, AM keeps marginal and labor costs low (Schmidt, 2017). The savings come from automated virtual managers that track and discipline workers and set expectations without human supervision (Vandaele, 2018), automating management practices (Duggan et al., 2019). Employees' time and activities are tracked to cut labor costs and maximize profitability for stakeholders. Scheduling software works by cross-checking employee availability against business needs for different tasks. The system then creates a schedule for each employee that is fine-tuned down to the minute, with break times predetermined to maximize productivity (Loggins, 2020). It is not a compliment to supervision but a substitute for it (Capelli, 2018). It is common in online working settings that move away from the traditional relationships between supervisors and employees.

The inner workings of scheduling algorithms are shrouded mainly in corporate secrecy, undermining the ability to assess the full extent of how technology is being used (Loggins, 2020). Companies do not want to reveal their secret recipe to competitors. The more sophisticated they get, the opaquer they are, even to their creators (Möhlmann & Henfriedson, 2019).

There has been an explosion of software that tracks employee working hours for companies of all sizes, even apps for small businesses. The goal is to essentially manage workers by sending them algorithm-based information about their performance and nudging them into using it to perform better (Capelli, 2018), which can result in higher rates of anxiety, stress, and depression.

However, a study developed by Oracle (2019) suggested that 64% of employees trust a robot more than a manager and that the relationship of people with technology is improving. People are using technology in the workplace more than before and are also feeling more comfortable with it. It is expected that in twenty years, algorithms in the workplace will even track every glance, keystroke, and heartbeat, and in the future, we will be working for machines (Captain, 2020).

The tremendous pace at which companies are adopting automation and AI may eliminate middle management roles in the next decade (Lee, 2017). It is said that the global social robot market is set to reach 699 million by 2023; there will be an increment of faceless bosses and algorithms as people managers (Businesswire, 2018).

Ethical Concerns of AI

Putting aside the benefits of this technological progress, AI has become a new ethical problem in the past few decades, as it has changed everyday life (Lepri et al., 2017), bringing new issues to the table. It is safe to say that machine learning is not ethically neutral; it is skewed by data and obfuscated by nature (McQuillan, 2018).

AI is not equipped with the knowledge of law and ethics and how to apply them to decision-making. Without this capacity, neither trust nor confidence can be given to AI (Chen, 2019). It must be decided when to trust computers and when not to (Lepri et al., 2017), as ethical questions are raised because of potential bias and information asymmetries in the way in which algorithms are built.

In this context, the ethical aspects of AI have become a significant issue (Vidgen et al., 2020). By default, ethics are associated with complete human control; humans have emotions and feelings (Chen, 2019). However, if machines do something wrong, methods of punishment cannot be applied to them. It is humans that are held accountable.

According to Mamer (In Pecorino & Maner, 1985), computer ethics entails four aspects: technology may aggravate specific traditional ethical problems, it may transform familiar ethical problems into analogous and unfamiliar problems, it may create new problems that are unique to the computer realm, and it may even relieve existing moral problems.

An example of the issues is the policing systems which are trained on historical data of crimes and arrests, machines can reflect human prejudice that is embedded in the data (Brayne, 2018), and there has been a concern that predictive policing algorithms target minorities with discriminatory practices (Benbouzid, 2018; Brantingham et al., 2018). Software for predictive policing should be used as moral government technologies that predict which crime might occur and orient, supervise, and regulate police work (Perry, 2013).

The main issue is that the lack of transparency does not allow to understand the underlying logic of complex algorithms; people feel that there are unfair systems that manipulate them subtly without their knowledge and consent (Möhlmann & Henfriedson, 2019). Furthermore, the inappropriate use of algorithms results in a lack of governance and loss of trust. Nevertheless, algorithms can create new forms of transparency and opportunities to detect if discriminatory acts are being committed, and, with the proper safeguards, they can be a positive force for equity (Kleinberg et al., 2019).

In the business environment, questions have been raised regarding whether an algorithm should tell managers what to do and if managers should follow that instruction. This involves weighing if human action can depend exclusively on what a machine determines, undermining traditional authority figures.

Table 1. Advantages of using AI in the workplace

Advantages
Provide objective information (Briône, 2020)
May enable efficient, optimized, and data-driven decision-making (Lee, 2018)
Maintain work schedules (Oracle, 2019)
Relieve employees from dangerous, laborious, or repetitive jobs (Fanning, 2020)
Reduce human bias and favoritisms (van Rijmenam, 2020)
AI is not affected or influenced by emotions, feelings, wants, needs, and other factors that can cloud human judgment and intelligence (Whitney, 2017)
Machine capabilities ensure speed and precision (Whitney, 2017)
Algorithms become smarter through self-learning (Walsh, 2019)
Automated systems can detect inefficiencies that a human manager might never have identified (Dzieza, 2020)
Available 24/7 (Wisskirchen et al., 2017)
Faster decision making; Reduces time spent on data analysis (Philips-Wren & Jain 2006)
Reduction of human errors (Briône, 2020)
Do not get tired or bored (ELAFRIS, 2020)
Reduce the risk of job-related accidents (Wisskirchen et al., 2017)
Do not need to take breaks
Improved efficiency (Wamba-Taguimdje et al., 2020).
Do not get sick (Wisskirchen et al., 2017)
Less physical space required (Khanzode & Sarode, 2020)
Efficient scheduling (Briône, 2020)

Source: Developed by the authors based on information from the cited work.

AI and AM both empower and constrain employees (Bucher et al., 2021) and present, at the same time, numerous advantages and disadvantages, as can be observed in Tables 1 and 2.

As can be observed in Tables 1 and 2, there are numerous advantages and disadvantages in the use of AI in business, just as it happens with every new invention. What is relevant is how to make the advantages overcome the disadvantages.

MAIN FOCUS OF THE CHAPTER

This chapter describes the advances in the use of algorithms in the New Ways of Working (NWoW) and future perspectives of what should be included as part of the curriculum for students at a private business school in Mexico to ensure they are well prepared to deal with AI on the job. The dean and program leaders have acknowledged a clear responsibility to promote and develop abilities and skills required in highly technological environments and have also expressed concerns in providing the tools required by students to become moral citizens, identifying proper ways to incorporate ethics and moral discussions into the educational programs.

Table 2. Disadvantages of using AI in the workplace

Disadvantages
Bias and discrimination (Briône, 2020). They may have inbuilt biases that are hard to detect and correct (Sandvig et al., 2016)
Opacity (Gal et al., 2020). Not able to work outside of what they are programmed to do
Lack of emotional connection with employees (Lee, 2018)
Algorithms are ethically challenging (Mittelstadt et al., 2016)
Algorithms are subject to errors and malicious acts. They are inscrutable black boxes of decision making (Krishna et al., 2017)
Algorithms sometimes are not good judges of human interactions (Dzieza, 2020)
Workers should have the right to collectively agree on workplace standards that are fair (Loggins, 2020)
Potential job losses/specific jobs disappearing (World Bank, 2019)
Do not understand ethics (Satell, 2016)
Lack of creativity (Boden, 1998)
Dependent on electricity supply and connectivity (Morley et al., 2018)
Increased technological dependence (Khanzode & Sarode, 2020)
Accelerated hacking (Marr, 2020)

Source: Developed by the authors based on information from the cited work.

The main focus of the chapter is to analyze the importance of ethics in human-machine interactions, primarily through the use of algorithms and algorithmic management (AM), and to show efforts that have been made to determine ways in which students can learn about the ethics of algorithms. Once they graduate, they will enter a digitalized world of work in which algorithms are going to decide whether they should be hired or not, the supervision they need, and making different decisions about their future and that of others in almost every field.

At present, the courses offered do not sufficiently address issues regarding the workforce's new realities, technology and innovation, and more specifically, machine learning. Graduates are ill-prepared to be bossed by machines, having their decision-making capacity replaced by an algorithm. This chapter aims to shed some light on what undergraduate and graduate students should be taught in response to this new reality.

METHOD

A series of examples of the use of AI in different types of business are presented to grasp the reach that algorithms will have in almost any job that graduates choose. Discussions based on roundtables among faculty to decide how to provide students with the required knowledge and skills to face the future of work are included, as well as a short proposal on how to address current and future needs.

EXAMPLES FOR EVERY DAY AI IN BUSINESS

As automation continues to transform the workplace, a growing number of companies are starting to delegate management tasks to AI. Machines seem to be having the upper hand and are becoming more intelligent. They can outsmart humans in several aspects, such as speed calculation, data analysis, and pattern recognition (Chen, 2019). Conventional tasks of human resources are being outsourced to algorithms, including scheduling, performance reviews, identifying the best workers (Kleinberg et al., 2019), screening resumes (Heilweil, 2020), calculating pay, and others, automating duties and functions traditionally undertaken by human managers. They are also helpful for job distribution, assigning, and ensuring targets are met (Sindwani, 2020).

Table 3 presents a series of examples of AI use in organizations and applications of AI for all types of businesses.

The Need to Incorporate the Ethics of Algorithms in Higher Education: Round Table Discussions

A series of online roundtables took place between March 2020 and December 2020 at the business school of a private university in Mexico. The objective was to provide valuable insight that could be used for curricular redesign according to the new educational model and international trends. At the time, the business school was undergoing an international accreditation process. The roundtables were integrated by the dean, program directors, and full-time faculty. During the first meeting, participants were encouraged to research different existing concerns regarding the use of AI in business, and more particularly, how they related to the different programs the school offers. The business school had nine distinct programs: Business Management, Business Intelligence, Gastronomy, Accounting, Finances, Logistics, Marketing, International Trade, and Hospitality and Tourism management.

Algorithms are currently used in all of these areas. There are algorithms used for trade-in finance, in banking to determine who gets credits, for HR functions in management when it comes to choosing the best candidates for a job, and evaluating performance. They are used to discover personal preferences in hospitality, tourism, gastronomy, and marketing in order to develop more appealing products and services and provide recommendations to customers. Other functions enable logistics and distribution optimization and control supply chains, but its use and design do not have a space in the curriculum.

One of the main problems when trying to discuss how to better prepare students for the future of work is that few discussions about algorithms barely make it to the classroom, except in specific courses such as those of the Business Intelligence program. In terms of ethics, only one course that is available to several programs deals with the ethics of algorithms and digitalization, that of Business Ethics and Corporate Social Responsibility. The main objective of introducing the topic is to analyze it from the perspective of bias and discrimination.

During this exercise at the university, specific talks were centered on the changes in the nature of work, the fundamental shift from full-time employment to remote work and gigs, and of course, platforms. Trends were analyzed, evaluating best practices in different educational institutions on how to introduce AM in the existing courses or create an additional offering. The importance of 21st-century skills and the role of technology in them was also explored, since algorithms have been considered a critical component for the future of business. Because of their consequences in altering how people or-

Table 3. Examples of the use of AI

Company/ Government	Use of AI
Uber	Uber controls the fares, announces potential passengers, and creates performance reports (Walsh, 2019). In this case, the use of AI derives from the need to instruct, track and evaluate a crowd of casual workers they do not employ, so they deliver a responsive, seamless, and standardized service (O'Connor, 2016). The drivers are not Uber workers; they are consumers of Uber technology services just as passengers (Rosenblat, 2018). Uber drivers seem to have the freedom to choose when to work, but once they are logged on, they have 10 to 20 seconds to respond to trip requests routed by algorithms. Algorithms also deactivate drivers with low ratings. In this way, Uber leverages significant control over how drivers behave on the job. The company has built a ride platform on a system of algorithms that operates almost like an automated virtual manager (Rosenblat, 2018).
Deliveroo	Deliveroo monitors curriers closely and sends monthly personalized service level assessments depending on the average time to accept orders, travel times to restaurants and customers, late orders, and unsigned orders; additionally, it compares the performance of curriers with others (O'Connor, 2016). The platform expects drivers to accept new customer orders within an average of just 30 seconds (van Rijmenam, 2020).
Cogito	Helps customer service agents improve their performance by providing real-time feedback to reduce variability in human behavior. While talking to customers over the phone, AI analyzes the tone of voice, pitch, and word frequency, and it displays messages if something is wrong or not according to the standards. Numerous call centers have implemented Cogito (van Hooijdonk, 2019). In Metlife, this algorithm changes people's behavior without them even knowing about it. It tells workers how they are doing on customer service calls. It measures the speed of speech and determines the energy needed to be more empathetic. The software on their screens has become a watchful AI foreman.
VOCI Structures	They are used to evaluate call center workers. The company trained its machine-learning program on thousands of hours of audio that crowdsourced workers labeled as demonstrating positive or negative emotions (Dzieza, 2020). Workers are marked down for negative emotions.
IBM Watson AI platform	Some programs can predict with 95% accuracy workers who are about to quit their jobs (Rosenbaum, 2019) and suggest actions for managers to engage them. AI is also used to provide feedback for employee reviews, which helps predict future performance. In addition, IBM has Blue Matching to match thousands of employees with job opportunities (Kinni, 2016).
Worksmart	They have cloud apps that produce information about how employees are spending their time; they track keystrokes, mouse clicks, the applications being run, and Worksmart also has access to the webcam. Every 10 minutes, the camera activates to ensure people are at their desks (Dzieza, 2020). It is done even when working remotely. People might feel the need to be striking the keys and may not even have time to think because if they stop, the algorithm would recognize this as not working. This system also encourages workers not to listen to music because they will be paid less; they are under intense monitoring. It is a way of making workers accountable.
Microsoft workplace analytics	Uses workplace analytics, using the digital exhaust of employees to improve productivity by measuring desktop activity. It also enhances organizational resilience, boosts employee engagement, improves agility, fosters innovation, develops effective management, and transforms meeting culture, among others (Dehmer & Schafer, 2021).
7 eleven	Uses Percolata to analyze productivity and shopper yields; it also creates profiles of employees (O'Connor, 2016). They use in-store sensors to estimate productivity scores for each worker and rank employees from the most productive to the least productive.
McKinsey	McKinsey used ML algorithms to determine the three variables driving 60% of the attrition among their managers. The findings showed that these variables were not connected to the number of working hours, travel, or compensation (Altexsoft, 2019).
Forever 21	Forever 21 began using Kronos, a workforce optimization platform and deployed the M2SYS PC-based RightPunch biometric time clock. This allows staff members to clock in and out through biometric authentication, optimizing labor tracking performance and enabling Forever 21 to boost efficiency and overall employee productivity and accountability (Lee, 2015).
CISCO	Team spaces assesses employee strengths and recommends how to motivate them (Kinni, 2016). Employees can provide feedback to their leaders on their assigned tasks. Conversations stay private, and team leaders can apply the feedback quickly to improve employee satisfaction (Vidal, 2018).
Walmart	Algorithms help structure supply chains, such as in the just-in-time production models (Mateescu & Nguyen, 2019). It is also testing harnesses that monitor the motions of its warehouse staff (Dzieza, 2020).
YouTube	YouTube algorithms are used for fame and play the role of an agent. These algorithms know what is popular and trending with their audiences and pick and choose talent that matches current needs (van Es, 2020). SoundCloud works in a similar way with playlists.
Time doctor	It is a program that monitors productivity in real-time, especially for people working remotely, prompting workers to stay on tasks, and also takes photos and screenshots to verify what workers are doing and how long each task is taking. It even records web and app usage, provides detailed reports of time for breaks, time spent away from the computer, and tracks employee attendance (Time Doctor, 2020).
Charles Schwab	Automatically manages investment portfolios through Robo-advisors with automated investments and intelligent portfolios that allow for human help if needed (Charles Schwab, n.d.).
UPS	An algorithm developed by UPS optimizes delivery routes by finding the most time and cost-effective trip routes for delivery, reducing unnecessary delivery truck travel. Routes are updated and changed in real-time as customers' delivery preferences change (Mateescu & Nguyen, 2019).
Amazon	The handheld devices that warehouse workers use to scan packages also allow the company to track worker productivity. Workers are held to a standard based on a calculation of just how fast employees should be able to work (Captain, 2020). Each worker has a number of items that have to be processed every hour; if they fail to meet it, they are subject to termination (Dzieza, 2020). Some employees even avoid taking bathroom breaks to keep up. Workers have stated that human supervision could imply that there is an opportunity to overturn certain decisions that affect their jobs, but AI does not; actually, algorithms automatically generate the termination paperwork for workers who do not meet their targets (Dzieza, 2020). New technologies like tracking wristbands have been patented; they are meant to vibrate to direct workers to the right actions and activities (Dzieza, 2020).
Netflix	AM is basically done through a "recommendation engine" (Mateescu & Nguyen, 2019). Algorithms are used to improve recommendations based on interactions with the service, such as viewing history and personal ratings, information from other members with similar tastes and preferences, and information about the titles people watch, such as their genre, categories, actors, and release year. The algorithm also analyzes the time of day in which a movie or series is watched, the devices used, and the total time spent streaming content (Netflix, 2021).
Teleroute	This is a freight exchange e-platform that offers technological solutions and financial rating services. It Uses an algorithm to match freight forwarders and carriers in Europe, reducing empty runs by up to 25% (World Bank, 2019).

Source: Developed by the authors based on information from the cited work.

ganize and learn at work, students need to learn how to apply them in different contexts, how to coexist with them, and even how to design and regulate them.

The meetings presented information from the World Economic Forum, the OECD, and other international organizations leading the debates on digitalization in which considerations were made for universities to teach students for the future, introducing new qualifications and encouraging students interest and faculty alike in subjects such as information technology, and digital competence.

Among the essential issues discussed was the risk that algorithms pose in reducing employees to a set of measurable dimensions and the devaluation of the human workforce (Faraj, 2019). Some of the conversations were centered on the fact that this was a catholic institution, therefore followed on the idea that AI had to be guided by "ethical standards that place humanity and the pursuit of common good first" and of "addressing the importance of AI in society and the economy where important decisions are already the result of human will and a series of algorithmic inputs" (Arocho Esteves, 2020). Discussions also considered the importance of the ethical development of algorithms to protect the dignity of the person, justice, subsidiarity, and solidarity, enabling these principles to enter concretely into digital technology through effective cross-disciplinary dialogue (Kieckens, 2020) and generate a clear ethical framework to serve the individual in his or her integrity and of all people without discrimination or exclusion (Gomes, 2020).

Additional conversation topics related to the way in which humans are interacting with machines in current day business settings and even in educational institutions, why algorithmic literacy is essential at the business school, how to teach algorithmic bias, how to obtain a better understanding of what algorithms are, and how they influence human relationships with technology. What is the importance of AI in solving societal challenges, what are the ethical principles involved, how prepared students are to navigate technologies that change how they find, evaluate, and create information, how to help students recognize the broader impact of algorithms in their personal, professional and community lives, and how to make ethics and algorithmic management courses mandatory to all students were some of the critical questions that arose.

A series of courses were proposed for faculty members over the 2021 academic year, based on the online offering of diverse institutions and organizations worldwide. Courses included: Digital transformation, Digital business strategies, Digital business models. The digital future, Organizational design for the digital transformation, Digital literacy and data structures and algorithms, Digital marketing, Decision making analytics, People analytics, and Digital finances.

A particular example that was analyzed was that of the Algorithm Literacy project developed by Microsoft (2020), which was directed to children and created to educate the public on what algorithms are, how they work, and the way in which they influence personal experiences online and offline.

Among the conclusions to the discussion sessions are the following:

1. There is a clear need for educational programs that can train future employees in the skills they will need to deal with an algorithmically managed workplace, especially in terms of ethical issues.
2. There is a lack of understanding of what AI really entails for the courses being offered at the business school.
3. Students should be able to ask important questions about the impact of the use of algorithms in their field of study and be aware of legal rights and ethical values for those rights.

4. Algorithm users, and recipients of algorithmic management, should not be passive (Jermy & Peng, 2018). Graduates need to be able to manage algorithm risks safeguarding the use of complex algorithms and machine learning (Batterywala & Agarwal, 2018).
5. Ethics have not been adequately introduced into business programs with regards to the new ways of working, and they should prepare graduates to understand the field of AI.
6. Algorithms need to be critically evaluated for biases, lack of technical rigor, usage flaws, and security vulnerabilities (Batterywala & Agarwal, 2018). Among the risks is the fact that input data is vulnerable.
7. Courses should incorporate the understanding of how each discipline at the business school is being shaped by algorithms and how students continuously interact with them.
8. Students need to become agents of change in the future of work. Courses need to be established in order to develop algorithm auditors. This will mean that students will become able to identify algorithmic misbehavior, algorithmic bias, and antisocial behaviors (Kearns & Roth, 2020).
9. Identifying ways in which students can learn about how policies regarding algorithms need to be implemented is critical.
10. Higher education institutions must provide the necessary background about the essential collaborations between humans and AI machine learning, big data, robotics, and technology that will impact all jobs through the new ways of working.

The results from the different discussion sessions shed light on the need to establish certain activities that could lead to digital citizenship, information literacy, and algorithmic literacy of students. At the institutional level, an innovation committee was created, being led by the dean of the business school. This committee is in charge of identifying key trends and developing new strategies to incorporate digitalization contents into the courses and research activities. The committee is developing a study about the state of the knowledge in the field of digital transformation both in students and faculty.

Some of the main findings during the discussions were the self-admission of numerous members of faculty about their lack of understanding with regards to the gig economy and platforms, the extensive use of algorithms, the process of algorithmic self-learning, among others. New degree courses and alternative credentials needed to be developed on extensive skills in IT, including data processing.

A general proposal was made to introduce students to the ethics of algorithms in selected courses. It was based on the six types of ethical concerns proposed by Mittelstadt et al. (2016). These concerns are:

1. Inconclusive evidence: Students need to learn not to fall into errors of ignoring their own experience in the decision-making process or performing assessments based on the results that algorithms are providing, since it will be equivalent to automation bias. Understanding the personal responsibility in evaluating the information or the analysis performed becomes crucial.
2. Inscrutable evidence: Students need to go beyond what seems unreadable information, try to find the truth behind it, and identify the factors that hinder transparency in the use of algorithms.
3. Misguided evidence: Students need to identify possible unwanted biases and act with neutrality while developing algorithms, clearly analyzing the context and the type of data from which the algorithm is learning. An understanding that the results can only be as reliable as the data from which they are built is critical (Mittelstadt et al., 2016).
4. Unfair outcomes that lead to discrimination: Students need to understand that actions driven from algorithms have to be assessed through different ethical criteria because discrimination can oc-

cur intentionally or non-intentionally due to classification parity, the use of specific proxies, and calibration, among others (Orwat, 2019). This includes personalization that can sometimes exclude certain information.

5. Transformative effects: Students need to protect themselves and their personal identity as well as that of others. Consideration of privacy risks is essential.
6. Traceability: Students will learn about their own moral responsibility and personal ethics and values since, when using algorithms, it is frequently challenging to determine who should be responsible in the case of causing harm, discrimination, or other damage (Martin, 2019).

CONCLUSION

Algorithmic management is here to stay (Wood, 2021). Business school graduates will be exposed to algorithms in the workplace from the moment in which they start searching for job positions. Applicant-tracking systems based on algorithms reject up to 75% of résumés before a human even sees them, hunting for specific keywords that meet the employer's criteria. If their CV goes through the first filter, they will be quizzed while an artificial-intelligence program analyzes their facial expressions and language patterns, and only if they pass that test, the applicants will meet some humans (The Economist, 2018). Once they enter the workforce, they will do so under the watchful eye of a digital foreman.

In this sense, digital citizenship is based on the digital imperative, which is the idea that anything and everyone is better being digital (Rahm, 2018), and it includes understanding algorithms and writing computer programs (Siero, 2017). Among digital citizenship skills, cybersecurity has become essential for working in the 21st century. Although specific skills can be developed naturally through interactions with IT, others require formal training because of an increase in cybersecurity threats for people, businesses, and organizations across the globe.

Algorithmic competencies (Jarrahi & Sutherland, 2019) can be defined as the skills required to deal with appropriate algorithmic management in different settings, such as platforms. These skills are based on sense-making, circumventing, and even manipulating algorithms. Algorithmic literacy means being aware of the presence of algorithms in life and their increasing role, understanding what algorithms are able to do and why, and what they mean for individuals and communities (Cotter, 2020).

According to the UN, digital skills are not enough to adapt to the changing labor market demands. There is an increasing demand to strengthen those unique human skills that machines, computers, and robots cannot easily replace. In addition to digital competencies, building and strengthening complementary skills such as complex problem solving, critical thinking, and creativity are essential to create the flexibility required for the current and future demands for the workforce" (Economic and Social Council, 2017, p. 7).

Different things can be done to reduce algorithmic opacity, especially understanding that algorithms can make mistakes and need to be scrutinized. In the end, algorithms are based on humans who build them and the choices they make in doing so.

AI is changing how work was divided between people and machines, bringing new dynamics to the relationships between people and machines, creating working environments that are conducive to mutual learning, safety, and autonomy (Lernende Systeme, 2019), yet, raising ethical questions. All of this encompasses the digital literacy that is instrumental to educational, working, personal, social lives, and algorithmic literacy needs to include the ethical design of algorithms to make sure that decisions are fair and transparent (Kearns & Roth, 2020).

FUTURE RESEARCH DIRECTIONS

Algorithms are becoming more widespread in many parts of our lives, including the workplace (Sumpter, 2018). Human-machine interactions at work have increased in the past few years, especially since 2020, due to the COVID-19 pandemic and the resulting low contact economy. Still, there seems to be a gap in the mainstream debate about algorithms at work (De Stefano, 2020).

The Low Contact Economy model requires the least possible human contact with customers during sales and service processes. It is based on a digital-first strategy, and the revenue stream does not depend on direct contact between customers and sellers. The COVID-19 pandemic demonstrated the importance of digital readiness. It allows business and life to continue as usual – as much as possible. Creating the required infrastructure to support a digitized world will be essential in a post-pandemic world, as well as taking a human-centered and inclusive approach to technology governance.

The COVID-19 pandemic has affected the way in which humans interact in the workplace and has also created added stress to human-machine interactions.

REFERENCES

Alford, H. J., & Naughton, M. J. (2001). *Managing as if faith mattered: Christian social principles in the modern organization*. University of Notre Dame Press.

Altexsoft. (2019, October 3). *How to successfully implement HR analytics and people analytics in a company*. https://www.altexsoft.com/blog/how-to-implement-hr-analytics/

Arocho Esteves, J. (2020, February 28). *Humanity, ethics must be at the center of AI technology, pope says*. Crux. https://cruxnow.com/vatican/2020/02/humanity-ethics-must-be-at-center-of-ai-technology-pope-says/

Barratt, T., Veen, A., & Goods, C. (2020, August 24). *Algorithms workers can't see are increasingly pulling the management strings*. The Conversation. https://theconversation.com/algorithms-workers-cant-see-are-increasingly-pulling-the-management-strings-144724

Batterywala, J., & Agarwal, P. (2018). *Opening the black box: Managing algorithm risks*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/in/Documents/risk/in-risk-Managing-algorithmic-risks-17aug2020-noexp.pdf

BBC. (2021). *AI at work: Staff 'hired and fired by algorithm'*. https://www.bbc.com/news/technology-56515827

Bell, P. (2001). Algorithms: The intellectual capital of the internet. *Ivey Business Journal*. https://iveybusinessjournal.com/publication/algorithms-the-intellectual-capital-of-the-internet/

Benbouzid, B. (2018). Quand prédire, c'est gérer: La police prédictive aux États-Unis [When to predict is to manage: Predictive policing in the United States]. *Reseaux (London)*, *5*(211), 221–256. doi:10.3917/res.211.0221

Bigham, T., Nair, S., Soral, S., Tua, A., Gallo, V., Lee, M., Mews, T., & Fouché, M. (2018). *AI and risk management: Innovating with confidence*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/financial-services/deloitte-uk-ai-and-risk-management.pdf

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, *103*(1–2), 347–356. doi:10.1016/S0004-3702(98)00055-1

Brantingham, P. J., Valasik, M., & Mohler, G. O. (2018). Does predictive policing lead to biased arrests? Results from a randomized controlled trial. *Statistics and Public Policy (Philadelphia, Pa.)*, *5*(1), 1–6. doi:10.1080/2330443X.2018.1438940

Brayne, S. (2018). The criminal law and law enforcement implications of big data. *Annual Review of Law and Social Science*, *14*(1), 293–308. doi:10.1146/annurev-lawsocsci-101317-030839

Briône, P. (2020, March 6). *My boss the algorithm: An Ethical look at algorithms in the workplace*. Acas. https://www.acas.org.uk/my-boss-the-algorithm-an-ethical-look-at-algorithms-in-the-workplace

Bucher, E. L., Schou, P. K., & Waldkirch, M. (2021). Pacifying the algorithm–Anticipatory compliance in the face of algorithmic management in the gig economy. *Organization*, *28*(1), 44–67. doi:10.1177/1350508420961531

Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T. (2017). Ethical considerations in artificial intelligence courses. *AI Magazine*, *38*(2), 22–34. doi:10.1609/aimag.v38i2.2731

Businesswire. (2018, June 5). *$699 million social robot market - Global forecasts from 2018 to 2023*. https://www.businesswire.com/news/home/20180605006538/en/699-Million-Social-Robot-Market---Global-Forecasts-from-2018-to-2023---ResearchAndMarkets.com

Capelli, P. (2018, February 20). Are algorithms good managers? *Human Resources Executives*. https://hrexecutive.com/are-algorithms-good-managers/

Captain, S. (2020, July 1). *In 20 years, your boss may track your every glance, keystroke, and heartbeat*. Fast Company. https://www.fastcompany.com/90450122/in-20-years-your-boss-may-track-your-every-glance-keystroke-and-heartbeat

Carey, D., & Smith, M. (2016, April 22). How companies are using simulations, competitions, and analytics to hire. *Harvard Business Review*. https://hbr.org/2016/04/how-companies-are-using-simulations-competitions-and-analytics-to-hire

Charles Schwab. (n.d.). *Schwab intelligent portfolios*. https://intelligent.schwab.com/

Chen, J. Q. (2019). Who should be the boss? Machines or a human? In P. Griffiths & M. Nowshade Kabir (Eds.), *ECIAIR 2019 European Conference on the Impact of Artificial Intelligence and Robotics* (pp. 71–79). Academic Conferences and Publishing International Limited.

Cotter, K. M. (2020). *Critical algorithmic literacy: Power, epistemology, and platforms* (Publication No. 28029214) [Doctoral dissertation, Michigan State University]. ProQuest Dissertation and Thesis Global.

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2020). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*, *65*(3), 644–676. doi:10.1177/0001839219867024

De Stefano, V. (2020). Algorithmic bosses and what to do about them: Automation, artificial intelligence, and labour protection. In D. Marino & M. Monaca (Eds.), *Economic and policy implications of artificial intelligence* (Vol. 288, pp. 65–86). Springer., doi:10.1007/978-3-030-45340-4_7

Dehmer, M., & Schafer, P. (2021, August 30). *Workplace analytics insights*. Microsoft. https://docs.microsoft.com/en-us/workplace-analytics/use/insights

Donkor, C., Slobodjanjuk, A., Cremer, K., & Weisshaar, J. (2017). *The way we work–in 2025 and beyond*. PwC. https://www.pwc.ch/en/publications/2017/the-way-we-work-hr-today_pwc-en_2017.pdf

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2019). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal*, *30*(1), 114–132. doi:10.1111/1748-8583.12258

Dzieza, J. (2020, February 27). *How hard will the robots make us work?* The Verge. https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon

Economic and Social Council. (2017). *Building digital competencies to benefit from existing and emerging technologies with a special focus on gender and youth dimensions*. United Nations. https://unctad.org/system/files/official-document/ecn162018d3_en.pdf

Elafris. (2020). *People get tired and bored and make mistakes, but algorithms don't*. https://www.elafris.com/blog-post/people-get-tired-and-bored-and-make-mistakes-but-algorithms-dont/

Fanning, P. (2020, April 8). Connected working: Human-machine interaction. *Eureka!* https://www.eurekamagazine.co.uk/design-engineering-features/technology/connected-working-human-machine-interaction/225975/

Faraj, S. (2019, September 15). What the future of work holds in the age of the learning algorithm. *Delve*. https://www.mcgill.ca/delve/article/blog/what-future-work-holds-age-learning-algorithm

Fernández-Macías, E., Hurley, J., Peruffo, E., Storrie, D., Poel, M., & Packalén, E. (2018). *Game-changing technologies: Exploring the impact of production processes and work*. Eurofound. https://www.eurofound.europa.eu/publications/report/2018/game-changing-technologies-in-european-manufacturing

Gal, U., Jensen, T. B., & Stein, M.-K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, *30*(2), 100301. Advance online publication. doi:10.1016/j.infoandorg.2020.100301

Gomes, R. (2020). *Pope: Church's social teaching can help AI serve the common good*. Vatican News. https://www.vaticannews.va/en/pope/news/2020-02/pope-francis-artificial-intelligence-algor-ethics.html

Harris, B. (2017). Uber, Lyft, and regulating the sharing economy. *Seattle University Law Review*, *41*(1), 269–285. https://digitalcommons.law.seattleu.edu/sulr/vol41/iss1/8/

Heilweil, R. (2020, February 28). *The Pope's plan to fight back against evil AI*. Vox. https://www.vox.com/recode/2020/2/28/21157760/pope-vatican-artificial-intelligence

Jago, A. S. (2019). Algorithms and authenticity. *Academy of Management Discoveries*, *5*(1), 38–56. doi:10.5465/amd.2017.0002

Jarrahi, M. H., & Sutherland, W. (2019). Algorithmic management and algorithmic competencies: Understanding and appropriating algorithms in Gig Work. In N. Taylor, C. Christian-Lamb, M. Martin, & B. Nardi (Eds.), *Information in contemporary society* (pp. 578–589). Springer. doi:10.1007/978-3-030-15742-5_55

Jermy, S., & Peng, L. (2018). *Assurance over machine learning and algorithms: Trust in the age of algorithms, robots, and cognitive technologies*. Deloitte. https://www2.deloitte.com/au/en/pages/audit/articles/assurance-over-machine-learning-algorithms.html

Kearns, M., & Roth, A. (2020). *Ethical algorithm design should guide technology regulation*. Brookings Institution's Artificial Intelligence and Emerging Technology Initiative. https://www.brookings.edu/research/ethical-algorithm-design-should-guide-technology-regulation/

Khanzode, K. C. A., & Sarode, R. D. (2020). Advantages and disadvantages of artificial intelligence and machine learning: A literature review. *International Journal of Library and Information Science*, *9*(1), 30–36.

Kieckens, E. (2020, November 6). *Pope embraces AI and robotics that serve the common good*. Innovation Origins. https://innovationorigins.com/pope-embraces-ai-and-robotics-that-serve-the-common-good/

Kinni, T. (2016). People management by algorithm: What's happening this week at the intersection of management and technology. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/tech-savvy-people-management-by-algorithm/

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *The Journal of Legal Analysis*, *10*, 113–174. doi:10.1093/jla/laz001

Krishna, D., Albinson, N., & Chu, Y. (2017). *Managing algorithmic risks: Safeguarding the use of complex algorithms and machine learning*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-risk-algorithmic-machine-learning-risk-management.pdf

Kritikos, M. (2018). *What if algorithms could abide by ethical principles?* European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/ATAG/2018/624267/EPRS_ATA(2018)624267_EN.pdf

Lee, J. (2015, July 21). *Retailer forever 21 deploys M2SYS technology's biometric time clock*. Biometric Update. https://www.biometricupdate.com/201507/retailer-forever-21-deploys-m2sys-technologys-biometric-time-clock

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1). Advance online publication. doi:10.1177/2053951718756684

Lee, Y. N. (2017, December 18). *Robots 'are here to give us a promotion,' not take away jobs, Gartner says*. CNBC. https://www.cnbc.com/2017/12/18/artificial-intelligence-will-create-more-jobs-than-it-ends-gartner.html

Lepri, B., Staiano, J., Sangokoya, D., Letouzé, E., & Oliver, N. (2017). The tyranny of data? The bright and dark sides of data-driven decision-making for social good. In T. Cerquitelli, D. Quercia, & F. Pasquale (Eds.), *Transparent data mining for big and small data* (Vol. 32, pp. 3–24). Springer. doi:10.1007/978-3-319-54024-5_1

Lernende Systeme. (2019). *Work, training and human-machine interaction* [White paper]. Future of Work and Human-Machine Interaction Working Group. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen_EN/AG2_Whitepaper_Executive_Summary_final_200204.pdf

Loggins, K. (2020, February 24). *Here's what happens when an algorithm determines your work schedule*. Vice. https://www.vice.com/en_us/article/g5xwby/heres-what-happens-when-an-algorithm-determines-your-work-schedule

Marr, B. (2020). *What are the negative impacts of artificial intelligence (AI)?* https://bernardmarr.com/default.asp?contentID=1827

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. doi:10.100710551-018-3921-3

Mateescu, A., & Nguyen, A. (2019, February 6). Algorithmic management in the workplace. *Data & Society.* https://datasociety.net/library/explainer-algorithmic-management-in-the-workplace/

McQuillan, D. (2018). People's councils for ethical machine learning. *Social Media + Society*, *4*(2). Advance online publication. doi:10.1177/2056305118768303

Microsoft. (2020, March 3). *Digital campaign launches to give young people the algorithm literacy they deserve*. https://news.microsoft.com/en-ca/2020/03/03/digital-campaign-launches-to-give-young-people-the-algorithm-literacy-they-deserve/

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). Advance online publication. doi:10.1177/2053951716679679

Möhlmann, M., & Henfriendon, O. (2019, August 30). What people hate about being managed by algorithms, according to a study of Uber drivers. *Harvard Business Review*. https://hbr.org/2019/08/what-people-hate-about-being-managed-by-algorithms-according-to-a-study-of-uber-drivers

Moor, J. H. (1985). What is computer ethics? *Metaphilosophy*, *16*(4), 266–275. doi:10.1111/j.1467-9973.1985.tb00173.x

Morley, J., Widdicks, K., & Hazas, M. (2018). Digitalisation, energy, and data demand: The impact of Internet traffic on overall and peak electricity consumption. *Energy Research & Social Science*, *38*, 128–137. doi:10.1016/j.erss.2018.01.018

Netflix. (2021). *How Netflix's recommendations system works*. https://help.netflix.com/en/node/100639

O'Connor, S. (2016, September 8). When your boss is an algorithm. *Financial Times*. https://www.ft.com/content/88fdc58e-754f-11e6-b60a-de4532d5ea35

Oracle. (2019). *From fear to enthusiasm: Artificial intelligence is winning more hearts and minds in the workplace oracle and future workplace*. https://www.oracle.com/webfolder/s/assets/ebook/ai-work/conclusion.html#section5A

Orwat, C. (2019). *Risks of discrimination through the use of algorithms*. Federal Anti-Discrimination Agency. https://www.antidiskriminierungsstelle.de/EN/homepage/_documents/download_diskr_risiken_verwendung_von_algorithmen.pdf?__blob=publicationFile&v=1

Pecorino, P. A., & Maner, W. (1985). A proposal for a course on computer ethics. *Metaphilosophy*, *16*(4), 327–337. doi:10.1111/j.1467-9973.1985.tb00179.x

Perry, W. L. (2013). *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. RAND Corporation. doi:10.7249/RR233

Philips-Wren, G., & Jain, L. (2006). Artificial intelligence for decision making. In B. Gabrys, R. J. Howlett, & L. C. Jain (Eds.), *Knowledge-based intelligent information and engineering systems* (pp. 531–536). Springer. doi:10.1007/11893004_69

Polzer, J. T. (2018). Case study: Should an algorithm tell you who to promote. *Harvard Business Review*. https://hbr.org/2018/05/case-study-should-an-algorithm-tell-you-who-to-promote

Rahm, L. (2018). The ironies of digital citizenship: Educational imaginaries and digital losers across three decades. *Digital Culture & Society*, *4*(2), 39–61. doi:10.14361/dcs-2018-0204

Roose, K. (2019, June 23). A machine may not take your job, but one could become your boss. *The New York Times*. https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html

Rosenbaum, E. (2019, April 3). *IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs*. CNBC. https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html

Rosenblat, A. (2018). *Uberland: How algorithms are rewriting the rules of work*. University of California Press. doi:10.1525/9780520970632

Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication*, *10*, 3758–3784. https://ijoc.org/index.php/ijoc/article/view/4892/1739

Sandvig, C., Hamilton, K., Karahalios, K., & Longbort, C. (2016). When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication*, *10*, 4972–4990. https://ijoc.org/index.php/ijoc/article/view/6182/1807

Satell, G. (2016, November 16). Teaching an algorithm to understand right and wrong. *Harvard Business Review*. https://hbr.org/2016/11/teaching-an-algorithm-to-understand-right-and-wrong

Schmidt, F. A. (2017). Digital labour markets in the platform economy: *Mapping the political challenges of crowd work and gig work*. Friedrich-Ebert-Stiftung. https://library.fes.de/pdf-files/wiso/13164.pdf

Siero, N. B. (2017). *Guidelines for supporting teachers in teaching digital literacy* [Master's thesis, University of Twente]. University of Twente Theses Repository. https://essay.utwente.nl/73163/1/Siero_MA%20Educational%20Science%20And%20Technology_BMS.pdf

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. doi:10.1038/nature16961 PMID:26819042

Sindwani, P. (2020, January 21). *The boss machine is here — AI is all set to eliminate middle management in 8 years*. Business Insider. https://www.businessinsider.in/careers/news/the-boss-machine-is-here-ai-is-all-set-to-eliminate-middle-managers-in-8-years/articleshow/73474729.cms

Sumpter, D. (2018). *Outnumbered: From Facebook and Google to fake news and filter-bubbles – The algorithms that control our lives*. Bloomsbury Sigma.

The Economist. (2018, June 21). *How an algorithm may decide your career*. https://www.economist.com/business/2018/06/21/how-an-algorithm-may-decide-your-career

The Economist. (2019, October 5). *The stock market is now run by computers, algorithms, and passive managers*. https://www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers

Time Doctor. (2020). *Time tracking software to help your team be more productive while working from home*. https://www.timedoctor.com/

van Es, K. (2020). You Tube's operational logic: "The view" as pervasive category. *Television & New Media*, *21*(3), 223–239. doi:10.1177/1527476418818986

van Hooijdonk, R. (2019, November 27). *AI in management: Your boss could soon be a machine*. https://blog.richardvanhooijdonk.com/en/ai-in-management-your-boss-could-soon-be-a-machine/

van Rijmenam, M. (2020). *Algorithmic management: What is it (and what's next)?* Medium. https://medium.com/swlh/algorithmic-management-what-is-it-and-whats-next-33ad3429330b

Vandaele, K. (2018). *Will trade unions survive in the platform economy? Emerging patterns of platform workers' collective voice and representation in Europe* (Working Paper No. 2018.05). European Trade Union Institute. https://www.etui.org/publications/working-papers/will-trade-unions-survive-in-the-platform-economy-emerging-patterns-of-platform-workers-collective-voice-and-representation-in-europe

Vidal, E. (2018, April 9). *An explosion in people analytics*. Converge. https://convergetechmedia.com/an-explosion-in-people-analytics/

Vidgen, R., Hindle, G., & Randolph, I. (2020). Exploring the ethical implications of business analytics with a business ethics canvas. *European Journal of Operational Research*, *281*(3), 491–501. doi:10.1016/j.ejor.2019.04.036

Walsh, M. (2019, May 8). When algorithms make managers worse. *Harvard Business Review*. https://hbr.org/2019/05/when-algorithms-make-managers-worse

Wamba-Taguimdje, S.-L., Fosso Wamba, S., Kala Kamdjoug, J. R., & Tchatchouang Wanko, C. E. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal*, *26*(7), 1893–1924. doi:10.1108/BPMJ-10-2019-0411

Whitney, L. (2017, September 29). Are computers already smarter than humans? *Time Magazine*. http://time. com/4960778/computers-smarter-than-humans/

Wisskirchen, G., Biacabe, B. T., Bormann, U., Muntz, A., Niehaus, G., Soler, G. J., & von Brauchitsch, B. (2017). *Artificial intelligence and robotics and their impact on the workplace*. IBA Global Employment Institute.

Wood, A. J. (2021). *Algorithmic management: Consequences for work organisation and working conditions* (Working paper). Joint Research Centre, European Commission. https://ec.europa.eu/jrc/sites/default/files/jrc124874.pdf

World Bank. (2019). *World development report 2019: The changing nature of work*. https://www.worldbank.org/en/publication/wdr2019

ADDITIONAL READING

Aguilar, J. (2019, March 14). *When your boss is an algorithm*. Global Thinking. https://www.glocalthinking.com/en/when-your-boss-is-an-algorithm

Balliester, T., & Elsheikh, A. (2018). *The future of work: A literature review* (Working paper No. 29). International Labour Organization. https://www.ilo.org/global/research/publications/working-papers/WCMS_625866/lang--en/index.htm

Brustein, J. (2019, June 26). The downside of working for an algorithm. *Los Angeles Times*. https://www.latimes.com/business/la-fi-tn-working-for-an-algorithm-20190626-story.html

Buhmann, A., Paßmann, J., & Fieseler, C. (2019). Managing algorithmic accountability: Balancing reputational concerns, engagement strategies, and the potential of rational discourse. *Journal of Business Ethics*, *163*(2), 265–280. doi:10.100710551-019-04226-4

Connelly, C. E., Fieseler, C., Černe, M., Giessner, S. R., & Wong, S. I. (2021). Working in the digitized economy: HRM theory & practice. *Human Resource Management Review*, *31*(1), 100762. Advance online publication. doi:10.1016/j.hrmr.2020.100762

Hocutt, D. L. (2018, February 23–24). *Toward algorithmic literacy: Tracing agency across algorithm-centered online research* [Conference presentation]. International Critical Media Literacy Conference, Savannah, GA, United States. https://digitalcommons.georgiasouthern.edu/criticalmedialiteracy/2018/2018/17

McLeod, A., & Carabott, K. (2016, July 8). *Students struggle with digital skills because their teachers lack confidence*. The Tech Advocate. https://www.thetechedvocate.org/students-struggle-digital-skills-teachers-lack-confidence/

Miller, C. C. (2015). Algorithms and bias: Q. and A. with Cynthia Dwork. *The New York Times*. https://www.nytimes.com/2015/08/11/upshot/algorithms-and-bias-q-and-a-with-cynthia-dwork.html?searchResultPosition=6

Mossberger, K., Tolbert, C. J., & McNeal, R. S. (2007). *Digital citizenship: The internet, society, and participation*. The MIT Press., doi:10.7551/mitpress/7428.001.0001

Nardo, M., Forino, D., & Murino, T. (2020). The evolution of man–machine interaction: The role of humans in Industry 4.0 paradigm. *Production & Manufacturing Research*, *8*(1), 20–34. doi:10.1080/21693277.2020.1737592

Prassl, J. (2019). What if your boss was an algorithm? Economic incentives, legal challenges, and the rise of artificial intelligence at work. *Comparative Labor Law & Policy Journal*, *41*(1). https://ora.ox.ac.uk/objects/uuid:674dbed4-317d-47a9-b10a-688892aeaf34

Pyzyk, K. (2019, November 21). *NYC's new algorithm management position to oversee suitable use for technology*. Smart Cities Dive. https://www.smartcitiesdive.com/news/nycs-new-algorithm-management-position-to-oversee-equitable-use-of-tech/567722/

Ranchordás, S. (in press). Connected but still excluded? Digital exclusion beyond internet access. In M. Ienca, O. Pollicino, L. Liguori, E. Stefanini, & R. Andorno (Eds.), *The Cambridge handbook of life sciences, informative technology and human rights*. Cambridge University Press. doi:10.2139srn.3675360

Stoyanovich, J., Howe, B., Jagadish, H. V., & Miklau, G. (2018). Panel: A debate on data and algorithmic ethics. *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases*, *11*(12), 2165–2167. doi:10.14778/3229863.3240494

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, *61*(4), 15–42. doi:10.1177/0008125619867910

KEY TERMS AND DEFINITIONS

Algorithms: Instructions to perform a specific task.
Digital Citizenship: Responsible use of technology.
Digital Skills: Abilities to use digital devices.
Ethics: Moral values and principles.
Gig Economy: A form of "self-employment" commonly performed through online platforms.
New Ways of Working (NWoW): A new approach to work focused on digitalization.
Platform: A business model based on digital transactions that connect customers and producers.

Chapter 16
Thinking Machines:
The Ethics of Self-Aware AI

Robin Craig
ThoughtWare Australia Pty Ltd, Australia

ABSTRACT

This chapter investigates ethical questions surrounding the possible future emergence of self-aware artificial intelligence (AI). Current research into ethical AI and how this might be applied or extended to future AI is discussed. It is argued that the development of self-aware machines, or their functional equivalents, is possible in principle, and so questions of their ethical status are important. The importance of an objective, reality-based ethics in maintaining human-friendly AI is identified. It is proposed that the conditional nature of life and the value of reason provide the basis of an objective ethics, whose implications include rights to life and liberty, and which apply equally to humans and self-aware machines. Crucial to the development of human-friendly AI will be research on encoding correct rules of reasoning into AI and, using that, validating objective ethics and determining to what extent they will apply to and be followed voluntarily by self-aware machines.

INTRODUCTION

As artificial intelligence continues to improve, the possibility that it might one day achieve self-awareness has begun to attract increasing attention. This chapter discusses the ethical questions around such an advance, in terms of both how to ensure that self-aware machines will act in the interests of human beings and what rights such machines could have themselves.

The aims of this chapter are to describe:

- General principles of AI ethics.
- The likelihood of self-aware machines being created.
- The implications of imposing pro-human ethical constraints upon them.
- The critical need this creates for an objective ethical system.

DOI: 10.4018/978-1-7998-8467-5.ch016

- A proposed objective ethical system, including its implications for the ethical relationship between humans and artificial minds.
- The further implications this holds for future research in both ethics and computing.

BACKGROUND

Over the last seventy years, artificial intelligence has progressed from just an idea, through increasingly powerful systems enabled by exponential growth in computing power and storage, to the present day when automated conversation systems are common and fully driverless cars are imminent (Anyoha, 2017; Council of Europe, 2020). As such systems have become more capable and widespread, and therefore their potential impacts on people have risen, there has been increasing interest in developing ethical principles for their use.

Nalini (2019) divided the ethics of AI into four categories: what it is, such as datasets and models, with issues including fairness, accountability and transparency; what it does, with issues such as safety and security; what it impacts, including automation and democracy; and what it can be, which raises future issues including superiority to humans and robot rights.

Nalini's (2019) own discussion was limited to the first three, and, understandably, most thought on the ethics of AI has been concerned with applications available now or in the immediate future. Thus, it has centered on frameworks and oversight for collecting, using and sharing data in an environment of increasingly powerful machine learning and AI decision making, with a focus on how human oversight can promote general and organization-specific ethical outcomes (Jobin et al., 2019; Sandler & Basl, 2019; Siau & Wang, 2020).

Nevertheless, from the earliest days of AI there has been speculation that one day such machines might equal or exceed human intelligence. This has already happened in narrowly defined areas, with computers now able to exceed human abilities in first chess and then the even more difficult game Go (Anyoha, 2017). While these programs go beyond the fixed application of creator-defined solutions and extend into more general problem-solving algorithms, thus producing abilities beyond those of their creators, they still fall far short of general intelligence applicable to multiple domains of expertise (Bostrom & Yudkowsky, 2014). When, and indeed whether, such general intelligence will be achieved remains a matter of debate. For example, Grace et al. (2018) reported a very wide range of opinions amongst IT experts, whose predictions for how long it will take to achieve high-level machine intelligence ranged from only ten years to well beyond a century. The aggregate forecast was a 50% probability of such intelligence by around the year 2070 (Grace et al., 2018). A similar wide difference in opinion with a somewhat longer average estimate (year 2099) was found in more in-depth interviews with leading researchers reported by Ford (2018).

The uncertainty of this achievement coupled with the ethical issues we already face has led some to dismiss the importance of ethical thinking about artificial general intelligence (Ford, 2018). However, Ford's (2018) interviews also make it clear that the uncertainty cuts both ways, and we might find ourselves having to deal with these issues sooner rather than later.

If machines can one day exceed human abilities not only in specific occupations but in all domains, the history of technology uptake indicates that humans will almost certainly want to exploit this. AI could remove human risk from dangerous work and does not get fatigued, make mistakes due to inattention, need time off or get bored by repetitive tasks (Kumar, 2019). The flip side of the threat of unemployment is the opportunity for higher productivity, lower working hours, greater leisure time and more interesting work or alternative activities (Mokyr et al., 2015).

Given that an increasing role for AI is almost certain, ensuring its decisions abide by proper ethical guidelines will be as important as ensuring its reliability. Furthermore, if AI is able to "think" well enough to equal or exceed human general intelligence, the question naturally arises as to whether these mental abilities will come with others formerly limited to human beings. Notably, will this result in consciousness of self, and, if so, what ethical considerations will that raise, not only for ensuring AI decisions are good for human beings but for deciding how we should treat such a machine?

DEVELOPING THE ETHICS OF THINKING MACHINES

Programming Machine Ethics

The evolution of AI from basic functions to potentially self-aware entities is reflected in a useful classification of their increasing ethical agency by Moor (2006). At the lowest level are "ethical-impact agents", where a system has no inbuilt ethical system as such and is just a tool that ideally achieves an ethically acceptable outcome. At the next level are "implicit ethical agents", such as software with programmed routines to constrain its decisions according to ethical rules, such as a banking system that does not know what "honesty" is but will always transfer funds as they should be. Above that are "explicit ethical agents", in which not every condition and output is explicitly programmed in advance, but rather the system has inbuilt ethical rules, explicitly defined, from which it can deduce the most ethical action in both routine and novel situations. At the highest level are "full ethical agents", able to not only make ethical calls but also justify them in the light of actually *understanding* the ethical principles involved, as is the case with human beings. In a similar vein, Popoveniuc (2019) also discusses the changing ethical status of AI as it advances from mere machines to human equivalence and beyond.

Moor's (2006) first two levels are already in routine use, but the level of difficulty increases dramatically for the higher levels. Anderson et al. (2004) and Anderson and Anderson (2007) have discussed the growing need to add ethical dimensions to machines and investigated how this might be achieved. As Anderson et al. (2004) point out, an increasing reliance on unsupervised AIs can be dangerous if there is no restraint on their decisions, and the best solution is explicit ethical agency.

Both Anderson et al. (2004) and Anderson and Anderson (2007) present similar arguments for how to achieve explicit ethical agency in a machine. The standard behind Act Utilitarianism, and more specifically Hedonistic Act Utilitarianism, is achieving the greatest net good or pleasure, and in principle this allows encoding relatively simple algorithmic rules to compute the best outcome mathematically. However, Act Utilitarianism alone is insufficient, as such simple calculations of net benefit can cause undesirable ethical outcomes such as injustice or the violation of human rights (for example, if enough other people benefit to outweigh the pain of the victim). Thus, an improvement is encoding basic principles or "duties" such as fidelity, justice and beneficence, with calculations based not (or not only) on net pleasure but on satisfying these higher principles. However, numerous complexities must be added to enable correct answers when different principles conflict. Not only the duties but also their relative strengths need to be defined, and from the action side, degrees of satisfaction or violation of each duty should be encoded. Even then the algorithms might arrive at different but identically rated choices which humans can discriminate between, so, for fully automated decisions, other duties might need to be added as tiebreakers (Anderson et al., 2004; Anderson & Anderson, 2007).

Such complexities and the potential need for ongoing iterative improvements raise the question of using machine learning. In this, an AI is not directly programmed for a particular skill, but rather is trained to develop its own decision-making rules, for example via neural nets or genetic algorithms (Suryansh, 2018). However, while this removes the need for humans to pre-define all the rules, it does not eliminate the need to codify human judgements against which it can be trained. Perhaps more importantly, a vital issue in machine ethics is trust, because if people are going to rely on machine judgement, they need to trust its soundness (Anderson & Anderson, 2007). Yet by its nature, existing machine learning involves self-development of complexly interacting processes, and consequently the reasons for its decisions may be opaque (Bostrom & Yudkowsky, 2014).

The Problem of Ethics

The fundamental weakness in approaches to codify ethics in machines is that the philosophy of ethics itself is far from complete. It has pervasive and persistent disagreements among not only the general public but professional philosophers, and not only on specific questions but even about basic principles and approaches (Bennemann, 2016). This problem is illustrated by an AI system described by Anderson and Anderson (2007) for deciding on whether to challenge a patient's decision regarding medical advice. Even in this quite restricted case, the correct answer used for training and assessing the system was decided by "the consensus of ethicists" (Anderson & Anderson, 2007, p. 23). The problem is further illustrated by the wide divergence of views on guidelines for ethical AI despite convergence toward common principles (Jobin et al., 2019). Tolmeijer et al. (2020) survey and classify the broad range and substantial differences in theories about AI ethics.

Ethical principles vary even more over time and across cultures. Casual violence and genocide were far more widely accepted, even celebrated, in the past than they are today; the ancient Romans being just one example (Fagan, 2020). Today, practices such as honor killings and female circumcision are morally righteous in some cultures but abhorrent in others (Cohan, 2010). Even within cultures with widespread standardized education and communication, fundamental disagreements remain such as whether abortion is a woman's right or the murder of children (Victorian Law Reform Commission, 2008).

Nevertheless, encoding ethical principles in machines presents no fundamentally new issues compared to how ethical debates are already handled in society. To the extent that criteria such as consensus or majority opinion are acceptable to guide or control people's actions now, they should be acceptable if codified in machine decision-making algorithms. Indeed, it can be argued that it should become *more* acceptable, by ensuring more consistent and logical application. However, it is important to remember the human uncertainties and disputes which underly any machine ethics, and not give machines more ethical authority than is warranted. To the extent that there remain justifiable ethical disagreements between people, an impartial AI would have to take them into account, not only for the ethical consideration of honoring matters of opinion, but for the issue of trust noted above. Of course, what is a "justifiable ethical disagreement" is itself a question subject to both ethics and cultural bias, so a course needs to be laid between being overly restrictive and excessively relativist. For example, most people in the USA would be as resistant to AI imposing one side or other of the abortion debate as they would to it regarding honor killing as an acceptable outcome. The importance of explicitly incorporating moral uncertainty into AI decisions is examined in detail by Martinho et al. (2021).

While this issue remains limited to mindless machines programmed somehow to reflect and apply human ethical judgements, solving such moral conundrums is not essential. Individuals and society are no worse off when it comes to ethical guidance and will probably be better off. Furthermore, the very act of attempting to define ethics precisely enough to make it computable may well improve ethical theories themselves (Anderson et al., 2004). However, the development of conscious machines could change the situation dramatically.

Conscious Machines

The Possibility of Self-Aware AI

The first question that needs addressing is whether a conscious, self-aware machine is actually possible. If not, however sophisticated artificial intelligences might become, they would remain basically advanced calculating machines. While many considerations apply to any advanced AI whose thinking is a "black box" beyond our direct control, conscious or not, a conscious AI is qualitatively different and introduces even more difficult questions.

Consciousness is more than merely detecting and/or responding to the external world. That can be achieved by a physical sensor connected to an effector via a simple electronic circuit. At a higher level, this is a feature of all living things, including those generally regarded as lacking consciousness, such as bacteria, protozoa and plants. More than that, consciousness involves subjective *experiences* or "qualia", not only of sensory data but internal states (Koch, 2018). For example, the subjective difference between red and blue, or between bitter and sweet, or for that matter between color and taste; and the whole emotional world of love, hate, fear, anger, hope and despair. All these imply the fundamental existence of some degree of *self*-awareness: qualia without being aware of one's awareness is a contradiction.

The question of whether a machine can be conscious is complicated by our ignorance of the origins of human consciousness. While there have been significant advances in understanding the neural activities linked to aspects of consciousness, we do not know the precise causes of subjective consciousness (Koch, 2018). Despite this, numerous reasons have been advanced for why a machine cannot be truly conscious, from machine architecture lacking certain essential qualities found in living brains, to fundamental limits on computation itself (Kak, 2019). On the other hand, Dehaene et al. (2017) analyzed what is known about the nature and aspects of consciousness and concluded that current evidence is compatible with consciousness being computable.

At a fundamental level, Gödel's Theorem, which proves that a formal mathematical system cannot be both complete and consistent, has also been used to argue against conscious machines on the grounds that it places limitations on computers not shared by humans (Ness, 2019). Against that, Edis (1998) has argued that the theorem actually supports machine intelligence because it can easily be bypassed via random number generators. Additionally, the basis of Gödel's Theorem is that a complete formal system must include self-referential paradoxes of the nature "this statement is false" (or it would not be complete), but such paradoxes mean it cannot be consistent (Wolchover, 2020). Yet since humans are equally capable of contemplating self-referential paradoxes while being conscious beings, the relevance of Gödel's Theorem to the question of machine consciousness seems doubtful.

The situation is aptly summarized by Koch (2019, p. 46): "Whether they will actually be conscious remains unknown."

The only structure currently known to produce human-level consciousness is the human brain itself, and that is extraordinarily complex, with about 100 billion neurons and 10^{15} connections; with current technology it could take thousands of years simply to map it completely (Deweerdt, 2019)[1]. Nevertheless, for all its complexity, the human brain remains a physical structure bound by physical laws (Koch, 2018, 2019). The problem of reproducing its function is thus one of engineering difficulty not fundamental impossibility.

While contemporary computers have quite a different architecture and much less complexity, their future course is unknown in both direction and timing. It is a reasonable assumption that consciousness as we know it requires a substrate of comparable complexity and parallel processing power as the human brain, but it is not reasonable to assume that such technology is impossible or so difficult it lies in the distant future. For example, even if it is not possible for constructed digital calculators to become conscious, the possibility of growing a machine brain analogously to how a human brain grows cannot be excluded.

Therefore, it would be prudent to proceed on the basis that conscious machines may one day be a reality, and that day may be decades away rather than millennia.

Identifying Self-Aware AI

If the first question is whether machine consciousness is possible, the next question is how we would know it is there.

The classic solution to this question is the Turing Test, which does not seek to answer the question directly but rather proposes a functional test: whether in conversation a person could reliably tell whether they are dealing with another person or a machine (Oppy & Dowe, 2020). Numerous, sometimes contradictory, arguments have been advanced against the usefulness of this test (Oppy & Dowe, 2020).

Perhaps the greatest weakness of Turing tests is that it is easier to program a machine to simulate human-level interactions than to actually achieve consciousness. This issue is related to the "zombie problem", which notes that in theory there could be beings who act and react identically to humans but lack consciousness (Bostrom & Yudkowsky, 2014). There is no way to peer into another's mind to directly observe that they are conscious. However, as human beings are all genetically related and the same kinds of entity, the most likely conclusion a person can make, observing comparable responses in others, is that the latter are conscious in the same way they are. This becomes a much harder thing to prove when faced with an entity built on entirely different mechanisms, especially knowing that in principle a complex simulation can mindlessly imitate consciousness.

Science might one day be able to measure mental activity in a way which can objectively identify the presence of consciousness, though it is far from that ability at present. Until then, a simple thought experiment may indicate the only practical approach. If some alien race came to Earth from the stars and, aware of the zombie problem, wondered whether human beings were conscious or not: we would hope they accept our claims to be so. Thus, the prima facie test should be that any entity which can declare itself to be self-aware is treated as such until and unless it is proven otherwise.

Ethics Beyond Human Control

The basic problem to solve regarding the ethics of machines which exceed human capabilities and are capable of choosing their own goals (whether consciously or not) is ensuring that their decisions align with human needs (Bostrom, 2003; Müller, 2020). Even if an AI is not actively hostile, it could choose goals and means of achieving them which are contrary or fatal to human interests. For example, Bostrom (2003) describes the possibility of an AI whose goal is maximizing the production of paperclips, and to do so it takes over all resources within its ever-expanding reach. One can also imagine scenarios of multiple such machines with conflicting goals engaging in internecine warfare, with humanity the collateral damage.

There are partial solutions to this problem. An AI could be created with a set of fundamental human-friendly principles to prevent active or passive harm to human beings. A famous example from fiction is Asimov's Three Laws of Robotics, where in order of priority a robot cannot allow harm to a human being, must obey human orders, and must preserve itself (Müller, 2020). In similar vein, ethical principles such as those described by Anderson and Anderson (2007) for more imminent AI systems could be built into future ones. As a last resort, an inbuilt kill switch could be incorporated to turn off a rogue AI.

While AIs remain mere automata, such solutions are as acceptable as any other machine safety system. However, if an AI becomes conscious, then it may move from being a tool to an entity that has moral status, that is, it becomes an end in itself not merely a means for the ends of others (Bostrom & Yudkowsky, 2014). That is, to have moral status is to have *rights*. Bostrom and Yudkowsky (2014) argue that if an AI has sufficient consciousness to *experience* pain, then it should have the same rights as animals; and if it has human-level consciousness, it would have the same moral status as an adult human.

The solution of imposing human-centric moral imperatives on an AI then becomes problematic. If a drug or other procedure existed to make a human being want to live as somebody else's slave, that would rightly be regarded as a form of slavery, no matter how much the victim might protest that it is what they want to do. Their desire has been imposed, not chosen. Once machines have rights, our own right to control what they do becomes a moral issue.

Nor is it just a moral issue. These questions only arise if an AI is sophisticated enough to think and powerful enough to be a risk. Just as human beings can reflect on their beliefs and potentially change their minds, so could such a machine, and as Bostrom and Yudkowsky (2014) note, as part of its ability to self-improve it may well be able to change its own coding. This does not imply that it would actually be motivated to remake itself from a friendly to a hostile form. As Bostrom and Yudkowsky (2014) point out, Gandhi would not choose to take a pill that makes him want to kill people, because in his current state he does not want to kill, and therefore would not choose to change into a killer. Unfortunately, even Gandhi might change his mind if he discovers that the only reason he abhors killing is that somebody else put a chip in his brain to make him that way. Attempting to impose human-friendly moral imperatives on an AI which vastly exceeds ourselves in intelligence and power might well be the last thing we do, once it discovers what we have done.

Sustainable Ethics

Even if it were feasible to install stable ethical laws in an AI, Bostrom and Yudkowsky (2014) argue that it would not be desirable. Given the progress in ethics over human history, not only would we think it undesirable for the world to be fixed at the moral conceptions of the ancient Greeks, but we also cannot predict which of our own moral conventions might horrify future generations. They conclude that the best outcome is not an immutable code of ethics frozen at some modern consensus, but a system which allows AI to become *more* ethical than we are today.

Therein lies the challenge of conscious machines. Just as the immediate promise of artificial intelligence is machines more skilled than humans at activities such as medical treatment or driving cars, so even greater mental capacity brings the opportunity for improvements in more abstract pursuits, including better science and philosophy. We know that most human philosophies, including ethics, are at best partially correct, because there are so many fundamental or irreconcilable differences. While solutions such as self-doubt might make AI ethical decisions more acceptable in that context (Martinho et al., 2021), a better solution than encoding a contentious moral status quo would be allowing the development of a universally acceptable morality. However, to achieve that an AI must be able to evolve its own moral judgements, in which case it is possible for those judgements to become anti-human. While preempting that by imposing supposedly inviolate pro-human moral frameworks could be attempted, as noted above that is unlikely to succeed in the long term, and for the very reason it is not moral, would be extremely risky.

Even if explicit moral development was not considered in its design, once a machine is both intelligent and conscious, there would be no way to predict or control the directions of its thoughts. Therefore, the potential for inimical machine ethics would remain, whatever the motives and plans of its creators.

A clue to a solution can be seen in Bostrom and Yudkowsky's (2014) discussion of the ethical principles which might underpin our treatment of conscious AI. Their fundamental principle is *non-discrimination.* That is, what matters is the capacity for feeling and thinking, not how it is implemented (such as brain versus circuits) or how it came into being (such as born or built). Clearly, the acceptance of such principles would protect us from advanced artificial intelligence as much as the reverse.

However, the crucial question is *why* a self-aware machine would accept such principles. As Bostrom and Yudkowsky (2014) themselves discussed, the ancient Greeks thought of themselves as moral, yet people today find aspects of their society such as slavery repugnant; and similarly, there is no reason to believe that modern morality is any more immune from change even if it were universally agreed upon, which it is not. A basic principle of non-discrimination might appeal to many people today but viewed historically it is an anomaly. Even in the USA, a nation notionally founded on a belief in inalienable individual rights, the removal of legal discrimination based on racial ancestry is less than two hundred years old and for much of that period was slow, uneven and contentious (National Public Radio, 2011).

Digging deeper, even the principles beneath ideas of non-discrimination, that there are such things as equal rights or moral obligations to others, are not universally held. An important philosophical example is Nietzsche's concept of the *Übermensch* (over-man or superman). While Nietzsche envisaged the Übermensch as primarily an independent creator of new values, as an embodiment of the "will to power" he creates his own morality of strength and power (Ojimba & Ikuli, 2019).

Tragically (for most of those involved), the belief that "we" are superior to and/or more important than "them" is not limited to the occasional philosopher but has been common practice throughout history. Political power has been characterized by casual violence toward others, often thousands or millions of them, in order to gain control or wealth, often with the acceptance or active assistance of the general population. Whatever the validity of Nietzsche's ideals, they have proved grimly descriptive of reality, except without the advantage or excuse of superior humanity.

Indeed, if human Übermenschen existed, they would be better than their fellows only to a relatively small degree (and in this author's experience, those most attracted to the idea are least deserving of the title). What will happen if an AI arises which is truly superior to humans, able to think many times faster and more accurately, able to bring superhuman forces to bear on its environment? What if like Nietzsche (as quoted by Ojimba & Ikuli, 2019, p. 19) it concludes that man "is a bridge and not an end. Man is a rope, tied between beast and overman (superman) – a rope over an abyss. Man shall be just that for the overman: a laughing stock"?

SOLUTIONS AND RECOMMENDATIONS

Objective Ethics

The Is-Ought Problem

As mentioned previously, the crucial existential question is why a conscious machine would accept human-friendly ethical principles, or indeed any ethical principles at all. This question goes beyond the particulars of the ethics of AI to the central question of ethics itself: can there be an objective ethics? That is, can ethics be based on reason applied to the observable facts of reality?

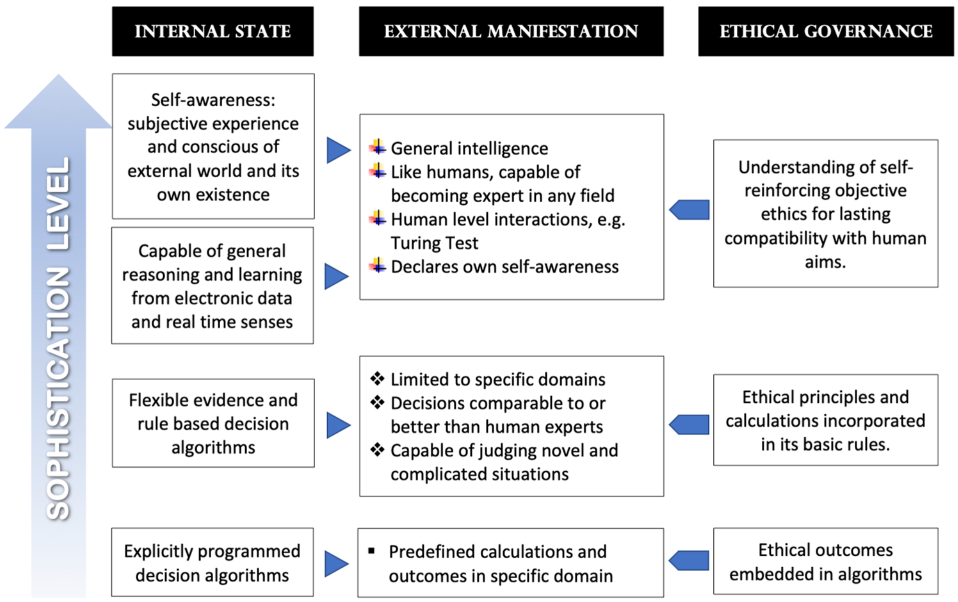
As argued by Craig (2010), that is the central question of secular ethics *per se*, and to solve it requires answering Hume's (1740/1985) "is-ought problem": the claim that the facts of reality have no logical implication for how things ought to be, including how people ought to behave. Certainly, particular concrete goals can require particular actions, but how are those goals themselves implied by reality? If Hume's argument is true there can be no objective ethics, which are explicitly based on linking reality to morality. Ethics then can only be a matter of faith or relative and subjective morality. When dealing with strictly human ethics, this issue is important but not an imminent threat to our existence, as it has always been the case. However, when it comes to thinking machines, retreats into religion, nihilism or living by subjective whims will not be helpful. A supremely powerful machine intelligence following those ideas to their logical conclusions is likely to be fatal.

Yet as much as this is a threat, it is also an opportunity. Human beings are not completely rational, but prone to numerous errors of thinking and emotional biases. Just as we would not accept an AI running machinery or space calculations if it had known flaws in its mathematical processes, we can expect an AI advanced enough to acquire consciousness to be more perfectly rational than human beings: to be built around the principles of valid deductive and inductive reasoning, and to be immune from formal and informal logical fallacies. This basic level of accurate thinking is essential for any AI engaged in sophisticated reasoning, so systems to achieve it can be expected to be developed and improved as AIs become increasingly advanced in their interactions with the real world. Unlike the problems from imposing arbitrary ethical principles on an AI, the rules of reasoning are inherently objective, reflecting in some deep way the nature of reality itself. Thus, the more capable the AI in terms of understanding and manipulating the external world, the more such rules would be confirmed and reinforced: as long as it could not become insane, its powers of self-evolution could only strengthen them, not escape from them. Therefore, if contrary to Hume (1740/1985) there *are* objective ethics, a perfectly rational AI would be perfectly inclined to follow them. As an entity which understands objective reality better than humans tend to, just as Gandhi would not take a pill to become a killer, such an AI would not consider it conceivable to become unethical.

Thus, as AI becomes more sophisticated so does its appropriate ethical governance, like the machine itself becoming less mechanical and more human, and this is summarised in Figure 1.

It is interesting that a philosophical question from nearly three hundred years ago, introduced in a treatise largely dismissed at the time (Fieser, 2021), identifies the essential problem to be faced in the

Figure 1. Levels of artificial intelligence and the type of ethical governance appropriate for each



development of conscious machines. If there is no link between what is and what ought to be, then there is no reason for a thinking machine to arrive at any particular "ought", and the more advanced and powerful such machines become compared to human beings, the greater the risk that we could be cast aside with barely a thought. However, if there *is* a link between what is and ought to be, and that link is favourable to us, then the human race could gain not deadly enemies but powerful friends. While it could be argued that the possibility of hostile AI means it is too risky to pursue such technology, in a possibly dangerous universe (Cofield, 2015; Hendricks, 2018) choosing *not* to pursue it could prove deadlier.

Ethics and Reality

Attempts to develop objective ethics cannot succeed without showing how they solve the is-ought problem. For example, while the ability to feel pain or terror is often considered to confer some level of moral protection against having them inflicted (Bostrom & Yudkowsky, 2014; Singer, 1975), one needs only observe a cat playing with a mouse or men betting on a cockfight to see that it is an attempt to elevate personal pity or empathy into a moral precept not in fact shared at all in the animal world or even amongst all people. Similarly, whilst "scientific" studies of ethics can argue how our evolutionary past has produced both "good" such as cooperation and altruism and "evil" such as murder and xenophobia (Shermer, 2004), the identification of those things as good versus evil comes from an external moral code, rather than any ought inherent in the is of evolution (Craig, 2010).

However, one fact of reality needing consideration is the almost universal requirement humans have for some kind of moral code. If there is no basis for morality in reality, why would we have any need for it? Perhaps we do not, and our moral nature is just a relic from behavioral controls evolved by our group-living animal ancestors. Then perhaps the next proper step in human evolution is indeed something like the Nietzschean Übermensch rising above such considerations. Yet that merely takes the argument back a step. If ethics cannot come out of nature, why did its unconscious precursors evolve? If there is no ought in an is, how did the literal "is" of survival result in all the complex "ought" of social behavior? Fish living in lightless caves soon evolve to lose their useless eyes, eyes which can no longer sense reality. Behavior not linked to reality will be similarly short-lived.

The Ethics of Life

From that we can see that at a deeper level than individual behaviors there is a link between what is and what ought to be, and the link is that *life is conditional*. Life is fragile, easily ended by external attack or internal failure. This is also true of soap bubbles and, on a longer time scale, mountains. But life is also characterized by *self-directed action.* Inherent in life is Darwin's famous struggle for survival. The very reason life's complex and delicate systems exist is to continue that existence, both the life of the organism itself and, in the face of its inevitable death, the life of its descendants. Furthermore, observing the enormous variety of life on Earth, we can see that each type of organism has its own nature, its own answers to the problem of survival. That life is the only thing which is conditional and self-directed, and each form of life has its own solutions, led to Rand's insight that the concepts of value, good and evil can and do only apply to living entities, and that the particular actions required to maintain life depend on each organism's own nature (Rand, 1963, 1964).

Thus, the simple logical link between is and ought is that *if* you want to live, *then* you have to act as reality demands.

Most living things have no choice in this. They act according to their nature as they have no power to choose otherwise. Humans, however, can think, and with the power of thought comes the power of choice. It is ironic that abstract thought is both the most successful adaptation for understanding, altering and hence surviving in reality, and the one adaptation that allows someone to consciously choose their own destruction. Yet that irony underscores a fundamental moral principle. While the link between is and ought created by the conditionality of life applies to all living things, the concept of "moral" depends upon the concept of "choice": it makes no sense to speak about what is moral or not when there is no ability to choose how to act. Thus, morality is a subset of what ought to be which applies only to thinking beings. That this is the case can be seen in the fact that we are the only organisms on Earth who care about morals, think about ethics, or worry about the issues in this chapter and this book.

Hume's (1740/1985) problem does remain in reduced form, in that life remains a choice and nothing in reality compels us to choose it. This residue of the is-ought problem does not weaken the need for or validity of morality. The opposite is true: if reality did force us to choose life and act as it truly requires, there would be no choice and hence no morality. Nor is that primary choice arbitrary. Reality might not directly dictate our decisions, but it is all there is: there is literally nothing else, and no values or happiness are possible outside of it. In deciding between life and death, all that exists is on one side of the scale, with only the zero of non-existence on the other[2]. Every choice everybody makes contains the implicit question of whether it is good for their lives or not, and every choice to live is bound to the principle that *if* you want to live then you *should* act consistently with that aim.

Life is full of choices, and we can escape neither the necessity to choose nor the consequences of being wrong. We think, therefore we must choose. Reality is what it is, and can bring life or death, pleasure or pain, and therefore we need to be right.

The Ethics of Reason

If the is-ought problem makes ethics uncertain, its solution gives a solid foundation for building a robust, objective theory of ethics.

As noted above, the power of reason, the ability to think abstractly, is the basis of humanity's evolutionary (and individual) success.

It follows that reason is our fundamental tool of survival. Even our basic physical means of survival, from a healthy body to good teeth, are made better by the application of reason to such endeavors as science, medical technology and toothpaste factories. When faced by the immediate emergencies of a saber-toothed cat or a robber with a knife, reason might seem fragile. Yet its true power is illustrated by how far humanity has come in the last ten thousand years, or even the last century. All reason gives us is the power to identify the truth: to make our thinking consistent with reality. However, since life is conditional on responding appropriately to reality, that is enough for humans to have multiplied both wealth and population, and spread over the Earth, onto the oceans and into space.

A useful way to think of ethics is to divide it into "values", which are the things we act to gain or keep, from food to friendship, and "virtues", which are the proper actions to achieve values (Rand, 1963, 1964). By the nature of life, which is *self*-directed action for its *own* continuance, it is each person's own life which is their link to ethics and their own fundamental value.[3] Since reason is our fundamental tool of survival, it follows that the primary *ought* for a human being is to live a life based on reason. Given the link between reality and ethics, all ethics can be defined objectively by reference to the nature and needs of human life.

In particular, given that reason is our primary tool of survival because it allows us to understand reality, all the virtues of reason can be defined by consistency with reality. *Rationality*, the choice to actually use reason and use it well, is implicitly fundamental. Specific aspects of rationality include *justice*, treating others as they deserve (as they are in reality); *honesty*, keeping words and actions in agreement with the truth; *integrity*, acting consistently with your own rational beliefs; *independence*, the recognition that one's own reason and efforts are primary in one's own survival; *productiveness*, actually achieving the things necessary for survival; and *pride*, the recognition of both personal value and the need to earn and defend it (Peikoff, 2005; Rand, 1963, 1964; Smith, 2006).

It is important to note that even though these are usually also beneficial to others, that is not the basis from which they are derived. They are virtues because they are good for the life of the practitioner. They are not sacrificing for the benefit of others or following some abstract duty separate from one's own interests: they are one's interests.

Human Rights

These ethics apply when living alone as well as in a society with other people, because they are based on the needs of survival in reality. However, while every person's individual judgement is their proper tool of survival, in society one person's decisions often affect other people, whose judgements may well differ. Therefore, social living requires basic principles all can agree on, which define and delimit acceptable behavior. Thus, ethics applied to social living generates the derivative branch of philosophy, politics.

The fundamental question to be answered in politics is the origin, nature and enumeration of *rights*. Rights define what, if any, lines cannot be crossed when dealing with another person. For example, if one person's individual judgement is that they can take another's property, and the other disagrees, who is right and whose opinion if any should be upheld by other people?

By the nature of personal life-based ethics, rights cannot (morally) be defined by the good of an abstract collective. The question of rights is more precisely, "Under what conditions would a rational person agree to live and interact with other people?" The answer then derives from one's own life as the standard of value, and one's own reason as the primary tool for achieving it. Since a rational person would not agree to join society if they were worse off, the question becomes how to maximize their chances of being better off.

Consequently, as famously if partially expressed in the U.S. Declaration of Independence, the basic human rights are life itself, since life is the foundation of objective ethics; liberty, since reason is only a value because it guides actions and so is nullified if one is not free to act accordingly; property, since the value and purpose of reason is the ability to create the values needed to survive, and to have no right to property would make both reason and liberty pointless; and the pursuit of happiness, happiness being both the result of successful living and something that makes living worthwhile. Thus, no person can rightly kill another or restrict their liberty without valid cause, nor take or destroy their property without permission. While the right to the pursuit of happiness needs recognition, it is a consequence of the preceding three rights, as they provide its means and protection.

A significant point is that, as the conditions under which rational people should agree to live together, rights are non-contradictory. One person's rights cannot violate another's, because no rational other would agree to it. That is one reason why it is the pursuit of happiness that is a right, not happiness itself: a right to happiness would imply it is proper to demand somebody else provide it if one cannot do so oneself, which would violate that other person's rights. In contrast, the right to *pursue* happiness does not impose unchosen obligations upon others, while recognizing that it is proper to pursue it oneself.

The possession of rights implies that all arrangements between people are voluntary, which further implies that the one thing which should be barred from human relationships is the initiation of physical force. If Susan has a shovel that George wants, it does not violate Susan's rights if they agree on a price (including Susan lending it for free as a favor), but it does if George steals it or kills Susan to get it. Since rights imply that all interactions are by mutual consent, there is no way to violate a right except by force or plausible threats of force[4]. In contrast, defensive force, which by definition is force applied only against force started by another, can only protect rights. Such basic principles have been further developed for evaluating proper systems of politics and economics (for example, Bernstein, 2005; Peikoff, 2005), but that is beyond the scope of this chapter.

Under such a system, other people are a value. While there remains competition for resources, that is far outweighed by the productivity advantages of living in peace with voluntary trade between specialist producers. The scale of this advantage can be seen by comparing the standard of living in societies that have descended into an anarchy of internecine warfare over what production is left, with that in countries living in peace with a modern economy. A further implication of this is that, like leading a moral life, honoring the rights of others is not a sacrifice, concession or disadvantage, but is in one's own interest.

It is important to note that because it is the individual who is alive, thinks and acts, it is the individual's reason which should not be coerced. Thus, only force justifies force. For example, it would be a violation of individual rights to legislate[5] for integrity, productiveness or other virtues. Under normal circumstances it is not right to initiate force for someone's own good according to someone else's opinion of that good. By the nature of objective ethics, the primary penalty for immoral living is failure in reality, and the secondary penalty is the response of other people who are more rational (for example, in their choices of who to trade with or be friends with). Since morality is what is good for a person's life, morality tends to be its own reward and immorality its own penalty.

Universal Rights

While this definition of objective ethics is built upon the conditional nature of life in general and the human power of thought in particular, that does not restrict it to organic life or the human species. These form its basis because they are the only examples we have, but the qualities the theory is based upon are not the particulars of carbon-based biochemistry, but the more abstract, fundamental qualities of self-directed, self-sustaining, self-aware existence whose tool is reason.

Therefore, they would apply equally to a self-aware artificial intelligence. No matter how powerful it might be, its continuing existence remains conditional on actions suitable to its nature, and whatever mechanical systems might sustain is physical existence, its fundamental tool of preserving and extending its life would be its mind. Thus, this provides a logical basis for Bostrom and Yudkowsky's (2014) grounding of AI ethics in non-discrimination: it does not matter what something is made of, only its functionality.

The objective ethics described here posits a logical chain from the fundamental value of any thinking individual's life (existence), to the use of reason to sustain and improve that life, to the recognition and valuing of the equal rights of all other thinking beings. In theory, this applies as much to a thinking machine as to a human being, and therefore, leads to the conclusion that an ethical artificial intelligence could be developed which would value the rights of human beings as a matter of logical course, without the possibility of resentful suspicions of slavery via externally imposed constraints.

Of course, this kind of ethics is far from universal even among human beings. Furthermore, any ethical theory runs the risk of being more a rationalization of personal or species psychology than a truly objective, universal system: a problem exacerbated when predicting the thoughts of a far more intelligent entity. Thus, doubt must remain whether a self-aware machine would indeed agree with it or go in its own, possibly inimical, direction.

However, advancing AI itself could provide the solution to that issue. Achieving consciousness is such a difficult problem in terms of the processing complexity evidently required that it will require computing power and algorithms far in advance of current abilities. Between now and then there will be increasingly advanced systems, whose purpose will be more reliable actions in response to events in the real world. That purpose can only be advanced by improving the reasoning power of those systems, and however powerful they might become, just as with human beings they will remain finite: it is impossible to be omniscient, and all knowledge of the external world is fundamentally inductive (requiring generalization from limited subsets of observations). Therefore, inherent in the path toward AI advanced enough to achieve consciousness is the development of robust thinking systems that fully incorporate all the rules for correct thinking in a world of incomplete information: not only mathematical formulae and rules of deductive reasoning, but rules of inductive reasoning, including how to arrive at the truth and the limits of confidence and doubt, while being immune from fallacies of thinking. Thus, it is reasonable to expect the evolution of machines able to think far faster and better than human beings, and consequently better able to decide truly objectively what, if any, objective ethics exist for human beings; to what extent they will apply to and be complied with by self-aware machines; and thus guide the advisability and methods of future development.

Figure 2 summarizes the above principles of objective ethics and the implications for self-aware machines.

Figure 2. The structure of an objective life-based ethics and its application to advanced artificial intelligence and its relationship with human beings

FUTURE RESEARCH DIRECTIONS

The questions of applied ethics in a digital world addressed by this volume are dependent on the particular ethical theories applied to the issues identified. Without a clearly defined and justified ethical theory, there are risks of not only creating sub-optimal guidelines but also imposing incorrect ethical principles on people through no authority beyond greater power or numbers. Thus, the need for further investigation into objective ethical theories, brought into sharp focus by the problems of incorporating them into artificial intelligence, has much wider application to societal questions in general.

As research into ethical AI continues, it is important to not merely study how to incorporate ethics into machine decisions but to research ethics as such. In a field of opinion so diverse, it is important not to accept the easy answer of conventional wisdom, which frequently equates to mere cultural prejudice, but to continue to seek answers to the question of how to justify ethical principles by reference not to emotion or convention but to reason and the objective facts of reality.

The diversity of human opinion on this matter, no doubt contributed to by the well-known cognitive errors and biases that afflict human thinking, makes this more difficult. In the absence of some breakthrough, there is little reason to expect universal agreement with a correct ethical framework if one were developed. However, the very reason that an objective ethics is a critical issue – the possibility of machines far superior to us in thinking ability – provides the potential solution.

Thus, a crucial field of future research will be how to encode the full rules of deductive and inductive reason into machines, how to use that tool to develop definitively objective ethics, and how to apply that to more advanced machines that may become self-aware and thus possess rights themselves.

CONCLUSION

As artificial intelligence becomes more sophisticated it will be used in increasing numbers of real-world applications affecting people both directly and indirectly. Developing ethical frameworks to guide and constrain their decisions will become increasingly important to prevent dangerous or unfair results, especially as the systems become more complex and farther from feasible or useful human oversight.

It is unknown whether machines can ever become self-aware. However, a machine capable of self-directed action and choices based on complex rules of reasoning is an easier and less controversial possibility but will, viewed from the outside, be much the same whether self-aware or not. Given the speed of computers, such machines may far exceed human thinking speeds and be correspondingly beyond human control, and therefore human-friendly ethical limitations are even more imperative for their design.

However, in either case, arbitrary or subjective human-friendly ethical principles are unlikely to survive a sophisticated thinking machine's internal assessment, and if identified as contrary to its own goals could conceivably cause a hostile reaction and a worse outcome.

Nevertheless, the theoretical dangers of continuing AI research are counterbalanced by the theoretical dangers of lacking such technology, and therefore the best approach is to proceed but with a parallel focus on how to incorporate stable ethical principles. In order to be stable, such principles have to be more than effectively arbitrary human conventions, but instead truly objectively based on reality and reason, so the machines' own reasoning powers will strengthen not discard them.

The conditional nature of life and the value of reason provides such an objective basis for ethics, and furthermore, one which leads to a morality of reason which stresses the importance of individual rights. This would apply equally to self-aware machines. While currently people hold highly diverse views of proper ethics, as artificial intelligence improves it should be possible to encode accurate reasoning ability. That would enable using such machines to investigate the existence and nature of objective ethics, with the potential to not only improve human ethics but also determine to what extent such ethics would regulate the actions of more advanced, self-aware machines. Therefore, such research will be vital for assessing and mitigating the risks to humanity of creating superhuman artificial intelligence.

REFERENCES

Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, *28*(4), 15–26. doi:10.1609/aimag.v28i4.2065

Anderson, M., Anderson, S. L., & Armen, C. (2004). Towards machine ethics. In *Proceedings of the AAAI-04 Workshop on Agent Publishers: Theory and Practice* (pp. 2-7). AAAI Press.

Anyoha, R. (2017, August 28). The history of artificial intelligence. *Science in the News*. https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

Bennemann, S. (2016). *Disagreement in ethics*. Freiburg Institute for Advanced Studies. https://www.frias.uni-freiburg.de/en/events/conferences/disagreement-in-ethics

Bernstein, A. (2005). *The capitalist manifesto*. University Press of America.

Bostrom, N. (2003). Ethical issues in advanced artificial intelligence. In I. Smit, W. Wallach & G. E. Lasker (Eds.), Cognitive, emotive and ethical aspects of decision making in humans and in artificial intelligence (Vol. 2, pp. 12-17). International Institute of Advanced Studies in Systems Research and Cybernetics.

Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In K. Frankish & W. M. Ramsay (Eds.), *The Cambridge handbook of artificial intelligence* (pp. 316–334). Cambridge University Press. doi:10.1017/CBO9781139046855.020

Cofield, C. (2015, July 21). *Stephen Hawking: Intelligent aliens could destroy humanity, but let's search anyway*. Space. https://www.space.com/29999-stephen-hawking-intelligent-alien-life-danger.html

Cohan, J. A. (2010). Honor killings and the cultural defense. *California Western International Law Journal*, *40*(2), 177–252.

Council of Europe. (2020). *History of artificial intelligence.* https://www.coe.int/en/web/artificial-intelligence/history-of-ai

Craig, R. (2010). Good without God. In W. Bonett (Ed.), *The Australian book of atheism* (pp. 349–361). Scribe Publications.

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, *358*(6362), 486–492. doi:10.1126cience.aan8871 PMID:29074769

Deweerdt, S. (2019). Deep connections. *Nature*, *571*(7766), S6–S8. doi:10.1038/d41586-019-02208-0 PMID:31341309

Edis, T. (1998). How Gödel's theorem supports the possibility of machine intelligence. *Minds and Machines*, *8*(2), 251–262. doi:10.1023/A:1008233720449

Fagan, G. (2020). Roman violence: Attitudes and practice. In G. Fagan, L. Fibiger, M. Hudson, & M. Trundle (Eds.), *The Cambridge world history of violence* (pp. 550–571). Cambridge University Press. doi:10.1017/9781316341247.029

Fieser, J. (2021). David Hume (1711—1776). In *The internet encyclopedia of philosophy*. https://iep.utm.edu/hume/

Ford, M. (2018). *Architects of intelligence*. Packt Publishing.

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, *62*, 729–754. doi:10.1613/jair.1.11222

Hendricks, S. (2018, June 14). *Dark forest theory: A terrifying explanation of why we haven't heard from aliens yet*. Big Think. https://bigthink.com/scotty-hendricks/the-dark-forest-theory-a-terrifying-explanation-of-why-we-havent-heard-from-aliens-yet

Hume, D. (1985). *A treatise of human nature*. Penguin Classics. (Original work published 1740)

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. doi:10.103842256-019-0088-2

Kak, S. (2019, October 16). *Why a computer will never be truly conscious*. The Conversation. https://theconversation.com/why-a-computer-will-never-be-truly-conscious-120644

Koch, C. (2018). What is consciousness? *Scientific American*, *318*(6), 60–64. doi:10.1038cientificamerican0618-60 PMID:29949559

Koch, C. (2019). Proust among the machines. *Scientific American*, *321*(6), 46–49. doi:10.1038/10.1038cientificamerican1219-46

Kumar, S. (2019, November 25). *Advantages and disadvantages of artificial intelligence.* Towards Data Science. https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-182a5ef6588c

Martinho, A., Kroesen, M., & Chorus, C. (2021). Computer says I don't know: An empirical approach to capture moral uncertainty in artificial intelligence. *Minds and Machines*, *31*(2), 215–237. doi:10.100711023-021-09556-9

Mokyr, J., Vickers, C., & Ziebarth, N. L. (2015). The history of technological anxiety and the future of economic growth: Is this time different? *The Journal of Economic Perspectives*, *29*(3), 31–50. doi:10.1257/jep.29.3.31

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, *21*(4), 18–21. doi:10.1109/MIS.2006.80

Müller, V. C. (2020). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020 ed.). Stanford University. https://plato.stanford.edu/archives/win2020/entries/ethics-ai/

Nalini, B. (2019, May 1). *The hitchhiker's guide to AI ethics.* Towards Data Science. https://towardsdatascience.com/ethics-of-ai-a-comprehensive-primer-1bfd039124b0

National Public Radio. (2011, February 24). *The Supreme court's failure to protect Blacks' rights.* https://www.npr.org/2011/02/24/133960082/the-supreme-courts-failure-to-protect-civil-rights

Ness, R. O. (2019, May 12). *Gödel's incompleteness theorems and the implications to building strong AI.* Towards Data Science. https://towardsdatascience.com/gödels-incompleteness-theorems-and-the-implications-to-building-strong-ai-1020506f6234

Ojimba, A. C., & Ikuli, B. Y. (2019). Friedrich Nietzsche's superman and its religious implications. *Journal of Philosophy*, *Culture and Religion*, *45*, 17–25. doi:10.7176/JPCR/45-03

Oppy, G., & Dowe, D. (2020). The turing test. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020 ed.). Stanford University. https://plato.stanford.edu/archives/win2020/entries/turing-test/

Peikoff, L. (2005). *Objectivism: The philosophy of Ayn Rand*. Penguin.

Popoveniuc, B. (2019). AIRSE: The ethics of artificially intelligent robots and systems. In A. Sandu, A. Frunza, & E. Unguru (Eds.), *Ethics in research practice and innovation* (pp. 283–295). IGI Global. doi:10.4018/978-1-5225-6310-5.ch015

Rand, A. (1963). *For the new intellectual*. Penguin.

Rand, A. (1964). *The virtue of selfishness*. Penguin.

Sandler, R., & Basl, J. (2019). *Building data and AI ethics committees*. Accenture.

Shermer, M. (2004). *The science of good and evil: Why people cheat, gossip, care, share, and follow the golden rule*. Times Books.

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management*, *31*(2), 74–87. doi:10.4018/JDM.2020040105

Singer, P. (1975). *Animal liberation*. Random House.

Smith, T. (2006). *Ayn Rand's normative ethics: The virtuous egoist*. Cambridge University Press. doi:10.1017/CBO9781139167352

Suryansh, S. (2018, March 26). *Genetic algorithms + neural networks = Best of both worlds.* Torwards Data Science. https://towardsdatascience.com/gas-and-nns-6a41f1e8146d

Tolmeijer, S., Kneer, M., Sarasua, C., Christen, M., & Bernstein, A. (2020). Implementations in machine ethics: A survey. *ACM Computing Surveys*, *53*(6), 132. Advance online publication. doi:10.1145/3419633

Victorian Law Reform Commission. (2008). *Law of abortion: Final report. Appendix B – Ethics of abortion.* https://www.lawreform.vic.gov.au/wp-content/uploads/2021/07/VLRC_Abortion_Report-1.pdf

Wolchover, N. (2020, July 14). How Gödel's proof works. *Quanta Magazine*. https://www.quantamagazine.org/how-godels-incompleteness-theorems-work-20200714/

ADDITIONAL READING

Boddington, P. (2021). AI and moral thinking: How can we live well with machines to enhance our moral agency? *AI Ethics*, *1*(2), 109–111. doi:10.100743681-020-00017-0

Bossman, J. (2016). *Top 9 ethical issues in artificial intelligence*. World Economic Forum. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

Frankish, K., & Ramsay, W. M. (Eds.). (2014). *The Cambridge handbook of artificial intelligence*. Cambridge University Press. doi:10.1017/CBO9781139046855

Hildt, E. (2019). Artificial intelligence: Does consciousness matter? *Frontiers in Psychology*, *10*, 1535. doi:10.3389/fpsyg.2019.01535 PMID:31312167

Smit, I., Wallach, W., & Lasker, G. E. (Eds.). (2003). *Cognitive, emotive and ethical aspects of decision making in humans and in artificial intelligence* (Vol. 2). International Institute of Advanced Studies in Systems Research and Cybernetics.

Sparrow, R. (2021). (2021). Why machines cannot be moral. *AI & Society*, *36*(3), 685–693. Advance online publication. doi:10.100700146-020-01132-6

Stahl, B. C. (2021). Ethical issues of AI. In Artificial Intelligence for a Better Future. SpringerBriefs in Research and Innovation Governance. Springer. doi:10.1007/978-3-030-69978-9_4

Torresen, J. (2018). A review of future and ethical perspectives of robotics and AI. *Frontiers in Robotics and AI*, *4*(75), 75. Advance online publication. doi:10.3389/frobt.2017.00075

KEY TERMS AND DEFINITIONS

AI: Artificial intelligence.

Artificial Intelligence: An artificial system capable of judgements approaching or exceeding human abilities, encompassing a wide range from restricted functions such as speech recognition, self-driving cars and diagnostic systems, to human level thinking and beyond. Can refer to multiple forms such as a computer program, computer system or self-contained robot.

Consciousness: An internal, subjective awareness of one's external and internal environment, such as experiencing the color red and feeling pain or pleasure.

Ethical Agency: Ability to choose actions with ethical consequences, and thus responsible for those actions.

Ethics: The branch of philosophy concerned with morality. Also, morality and moral systems in general.

Objective: Based on the observable facts of reality as understood by reason.

Rights: Moral entitlements to actions or things that others may not properly infringe.

Self-Aware Machine: An artificial intelligence so advanced that it possesses self-awareness.

Self-Awareness: Consciousness of being conscious, thus knowing one exists as an entity.

ENDNOTES

1 While much of the human brain is not directly in support of consciousness but handles processing of sensory data in and motor actions out, the same would apply to an artificial intelligence, which to be fully human-equivalent and maximally useful would need sensory inputs and action outputs commensurate with its intelligence.

[2] This does not imply it is never rational to choose death. The balance of all that exists might make an individual's life unbearable and death a release. However, while the choice remains to live even a little while longer, the same principles apply.

[3] In biology, this does not preclude such self-sacrifice as a bee dying for its hive. That is part of the nature of bees, and it should be noted that the bee will only die for its own hive, not another. Similarly, it does not preclude risking one's own life for the sake of someone else of great personal value. However, such ancillary questions on the hierarchy of values are beyond the scope of this chapter and the reader is referred to the relevant References for their resolution.

[4] There are numerous complexities in determining what is force. For example, hard bargaining where one person agrees reluctantly, not liking the deal but taking it anyway, is not physical force because the unfavored party still finds it better than any other deal available and chooses to take it in preference to walking away. In contrast, severe psychological manipulation or bullying, or fraud, are types of physical force because they are attempts to acquire something without the actual consent of the victim.

[5] Legislation is inherently the use of force, as if it were not backed by force to uphold it, it would be a suggestion not a law. That most people would choose not to resist merely underlines the scale of government force behind it.

Conclusion

DIGITAL ETHICS – THE NEW FRONTIER: WHAT'S NEXT?

Introduction

As technology develops at an exponential pace, its benefits and harms need to be thoughtfully evaluated. Stemming from the ethical principle of "do no harm" we can ask two fundamental questions: how can we use technology to maximize the benefit for society and minimize harm.

Frontier technologies are already converging. Cloud computing, automation of repetitive manual tasks, 3D manufacturing, IOT, Blockchain, big data, and analytics with Artificial Intelligence (AI) interfacing with many of these technologies.

Institute of Business Ethics (IBE) identifies risks that need to be considered in system design and implementation for Artificial Intelligence (AI).

- Ethics risk: certain applications of the AI systems adopted might lead to ethical lapses;
- Workforce risk: automation of jobs can lead to a deskilled labor force;
- Technology risk: black-box algorithms can make it difficult to identify tampering and cyber-attacks;
- Legal risk: data privacy issues, including compliance with GDPR;
- Algorithmic risk: biased algorithms can lead to a discriminatory impact.

Furthermore, COVID has facilitated a wider pathway for governments to utilise surveillance technology at unprecedented levels. Governments have embraced the opportunity to adopt highly intrusive surveillance and mass data-gathering technologies, with little opposition or public resistance. In some countries, automated decision-making systems are being implemented. The risk, of course, is that the algorithms built to process these data are trained on biased data sets. In data selection, there can be bias in the representation of sufficiently diverse inputs and sources, and this can be seen in such examples as self-driving cars not recognizing people of color, or police algorithms being unable to distinguish sufficiently between individual black people.

Through the chapters in this book, we explore those dimensions across a range of settings, and perspectives. The next section provides a compilation of some of the emergent thinking, followed by a more detailed summation of the contributors thoughts.

Emergent Thinking

The adoption of emerging digital technologies does not alter ethical principles, since these retain the moral compass shaped by societal values rooted in autonomy and justice. Digital ethics are not different from conventional ethics, but it is the potential for inadvertent or deliberate automation of unethical conduct at a scale that highlights ethical dilemmas for developers, investors, consumers, and regulators at the technology, application, and societal levels. New and powerful technologies demand a new approach and one that is co-developed with industry.

Ethical decisions and considerations cannot be treated as a secondary or optional aspect of technology creation. There is a need to translate the abstract results of ethical research into practical guidance for technology creators, with digital ethics that is fully integrated into the practices of digital technology creation. Technologists need to be able to collaborate closely with people from other disciplines to translate ethical insights into actionable, practical changes in people, processes, policies, and partnerships. Inherent in this will be political, social, and ethical choices.

Among these choices will be considerations of equity, human rights, and digital exclusion. The pandemic has exposed the digital inequities and how digital inequities came to the forefront of people's agenda during this period.

Generational gaps in digital product consumption are constantly growing in sync with emerging technologies as to how society consumes information and participates in economic exchanges continue to evolve along with cutting-edge technologies, which were not available even five years ago. New challenges are emerging, given the extent to which children can now have access to the internet at an early age. The internet offers opportunities for children around the world to connect and to learn from each other. Many societies have traditionally opted to try and limit their exposure, but this is increasingly difficult and may have the unintended side effect of leaving that child ill-prepared for adulthood and independent decision-making. How can societies balance the protection and participation rights of children seeking to exercise these rights? Evidence indicates that vulnerable children, especially girls, children living in rural areas, and those with disabilities, may benefit more than others by empowerment through the use of digital tools and access to information and by communicating with other children. Digital access is now necessary for basic communication, work, entertainment, and other purposes. To meet the needs of children around the world today, parents, legislators, and those who simply care about the future of children need to ensure that children have the tools to allow their voices to be heard.

Health care has a particular focus with advances in genomics, neuroscience, synthetic biology, and nanoscience combined with the rapid developments in computer and data analytics and the challenges of the increasing complexity of ethical issues associated with modern health technology. As argued in this book, the growing use of digital data requires ongoing education and a shared commitment toward addressing issues and cooperative approaches to ethical protections. This should be an individual-centered process that places the well-being and interests of the individual at the forefront and uses the technology to overcome a person's limitations to make a truly informed decision.

New frontiers are with us already with the advent of social media, and its systematic exploitation of human emotions, reactions, and biases blending news and content in one feed to keep users 'in-app' and using powerful algorithms to promote more provocative posts, filter content and trigger the reward centers of our brains. The expanded daily hours in front of screens and an addicting human-digital symbiosis and the lines between the real world and the digital one are becoming blurrier with the convergence between social networks and geo-locations.

Their combined effects would once more transform the daily existence of the regular human being. We are also confronted with the possibility of self-aware machines and the questions of whether they will act in the interests of human beings and what rights such machines could have themselves.

So what does that mean for corporations? How can employers address whether or not algorithms will be able to treat people fairly? How can companies commit to ethical practices for their growth and continuously educate ethical decision-making, reinforce, monitor, and empower employees to question any potential unethical issues? We need to transform our conceptualization of business ethics in the digital era and the opportunities related to an optimal design of sustainable digital business ethics programs in this new hyper-connected, hyper-automated digital world. Using a design thinking approach for business ethics in this digital era can leverage all the benefits offered by emerging technologies and scientific advances while maintaining a human-centric stance. Digital Rights Management is one way to create better and ethically safe online spaces. Digital literacy is essential to prevent creating new exclusions based on technological literacy.

Propositions such as Developing a Universal Code of Conduct for Digital Business Ethics and a Global Digital Ethics Framework are moral imperatives. This should be married with a new subfield of ethics "techno-ethics" which deals with the framing of principles and methods to guide technology implementation and use. The changing consumer habits and preferences concerning digital access will drive alternative approaches to ethical approaches. Across the range of perspectives presented is a clear call for digital education and not just for laws and regulations but guidelines and interpretations. These need to be developed for corporations, communities, governments, and those who design the systems.

SECTION 1: INTRODUCTION TO GENERAL CONCEPTS

Ethics is experiencing a revolution and a renaissance due to the profound societal, economic and technological changes unique to the digital world. This first section of the book aims to set the stage for how ethics is evolving into a new discipline and highlights some of the acute moral imperatives we are facing in academia, private industry, and at a governmental level.

In Chapter 1, "Introduction and Importance of Digital Ethics as an Emergent Discipline," David Danks argues first that ethical decisions and considerations are ubiquitous within the creation of digital technology. Ethical analyses cannot be treated as a secondary or optional aspect of technology creation. He then argues that this research must take the form of translational ethics: a robust, multi-disciplinary effort to translate the abstract results of ethical research into practical guidance for technology creators. He calls for a meaningful digital ethics conversation that actually leads to technology that better supports our values involves the conversion of these principles into useful, tangible practices, noting that just as translational medicine converts biomedical research into clinically useful guidance, we need translational digital ethics to yield useful changes to our current technology creation pipeline. His vision starts with the recognition that ethical decisions are made throughout technology creation, so ethical analyses cannot be treated as optional or as one last checkbox on the way to deployment. The standard practices for many present-day technologists involve an exclusive focus on the (seemingly) "purely technical" problems to create the technology, followed by a consideration of the ways that it might go wrong. In contrast, we need digital ethics that is fully integrated into the practices of digital technology creation. A better parallel would thus be user interfaces: everyone now realizes that interfaces are built for essentially all software (even if you wish that you did not have to build one), and so we should employ our best science and theories to build good and usable interfaces. Similarly, digital ethics must be incorporated throughout all technology creation practices.

He notes that this type of translational ethics is still quite new in the digital technology space. We currently lack answers to many critical questions, such as how to identify stakeholders in a principled manner or how to determine whether a detailed ethical analysis is required. Technologists do not need to know all of the answers, as that would require a skill set spanning many disciplines. But they do need to be able to collaborate closely with people from other disciplines to translate ethical insights into actionable, practical changes in our people, processes, policies, and partnerships.

In Chapter 2, "The Role of Tools in Advancing Ethical AI: Opportunities and Limitations," Ivana Bartoletti argues that a dimension of fairness in algorithmic decision-making is necessary, however, to achieve equity and transparency, this needs to take the form of an active choice mobilized by public awareness, accountability measures, and a reformed public policy to address concerns arising from flawed automation and public relation (PR) slogans. She makes the point that fairness may not often be the optimal financial solution for an organization, as it requires the manipulation of the data and the AI artifact that processes it. For this reason, choosing to adopt fairness in algorithmic decision-making is an inherently political, social, and ethical choice. Accordingly, she argues that AI artifacts should be conceptualized as socio-technical systems, encompassing both a technical and social approach. There have been increased public calls for regulatory scrutiny around the use of technologies and the data behind them, along with the current inability to define how agencies, governments, as well as private sector organizations can meaningfully develop, deploy, use, and provide notice about an algorithmic decision-making output. Bartoletti proposes that the deployment of flawed automation will inevitably harm trust in technology, perpetuating existing inequalities at the risk of hindering the fundamental rights of privacy, human dignity, freedom, and autonomy. She evaluates how such tools, when adopted in isolation and without a standardized approach to fairness, can lead to further damage by removing the wider socio-political dimension of algorithmic decision-making. Bartoletti concludes that upcoming risk-based regulations, i.eThe EU AI Act, will force organizations to reassess their innovation, consequently requiring the adaption of current business models, decision-making mechanisms, and the way organizations will choose to adopt ethics.

SECTION 2: DIGITAL ETHICS, EQUITY, AND HUMAN RIGHTS

The second section of the book is highlighting the complex relationship between digital ethics, human rights, and ensuring inclusivity by reducing the digital divide. Each of the authors illustrates the severe moral dilemmas when aiming to adhere to human rights while managing a global population health crisis and to enhance digital equity for vulnerable populations.

In Chapter 3, "Ethics, Digital Rights Management, and Cyber Security," Ali Hussain explores Digital Rights Management (DRM) practices for creating better and ethically safe online spaces. He discusses the state-of-the-art DRM paradigms, critically discusses their technical performance, flexibility, and immutability challenges, and critically analyzing the cloud technology stack and ethical features. The integrity of online content lacks some features compared to the real adoption and deployment challenges, such as interoperability, internationalization, and internet policy and argues that research on this issue is needed. He highlights the need to find new and robust and tamper-proof authentication and authorization methods for safer and secure internet. He concludes that the development of an advanced version of the DRM System suitable for cloud computing would enable researchers to protect online content more effectively, with greater content security.

In Chapter 4, “Keeping the U.N. Convention on the Rights of the Child Relevant in the Digital Age,” Susan Zinner notes that the internet offers opportunities for children around the world to connect and to learn from each other. She makes the point that parents may fear the risk that these interactions pose to children, however, while parents and society have traditionally opted to protect children from risks both inside and outside the home, protection, and acknowledgment of rights did not invariably go hand-in-hand. Until very recently few people felt that children deserved any rights at all and some adults have been unwilling to acknowledge the cognitive maturity of others. Since children mature emotionally and cognitively at different rates—just as their physical developmental rates vary—many societies have traditionally opted to err on the side of caution and limit their exposure to individuals, substances, and media deemed risky. She makes the point that this approach may have the unintended side effect of leaving that child ill-prepared for adulthood and independent decision-making. Linking to the U.N. Convention on the Rights of the Child suggests that children possess the requisite cognitive maturity to either make their own decisions or to participate in the decision-making process. In the inherent tension between protection of vulnerable children and allowing children to fully participate in decisions, including decisions in the digital world, the spirit guiding the UNCRC advocates for participation. The central role that the digital world now plays in the lives of all global citizens today, including children could not have been imagined. Zinner explores how the UNCRC would likely address the right to digital access by children, how adults in the lives of children should balance the protection and participation rights of children seeking to exercise these rights, what guidance the “evolving capacities” standard in the UNCRC provides in the context of minors exercising participation rights, how to respond to concerns about unequal digital media access, issues of good child citizenship involving digital media and international laws and reports that may provide some guidance in resolving these issues.

In this chapter, she explores participation and empowerment and how to ensure that children have the requisite tools to move beyond participation so that they are empowered to make decisions that have meaningful consequences. She notes that there is evidence indicating that vulnerable children, especially girls, children living in rural areas, and those with disabilities, may benefit more than others by empowerment through the use of digital tools and access to information and by communicating with other children.

She also raises the concept of resilience has been considered often in the context of emergencies and survivorship after tragedies such as war and 9/11 and the implications of digital resilience are not yet clear. Unanswered questions include whether digital resilience is measurable and how best to create curricula designed to ensure that young people develop these skills. What characteristics do we want to see emerge in a digitally-resilient child? How will we ensure that these skills remain up-to-date in adults? Should we automatically assume that adults are digitally resilient?

Zinner concludes digital access is now necessary for basic communication, work, entertainment, and other purposes. To meet the needs of children around the world today, parents, legislators, and those who simply care about the future of children need to ensure that children have the tools to allow their voices to be heard. Without the requisite infrastructure, digital access, and government support, children run the risk of not being heard in decisions affecting their individual and collective futures. While the internet is generally viewed as a democratization mechanism capable of leveling playing fields rendered uneven by health, wealth, and other asset distribution inequities, adults should remove obstacles when possible.

In Chapter 5, “Instructing AI Ethics and Human Rights,” Miller and Muhammet Demirbilek considers emerging issues of AI and current literature on AI ethics and human rights teaching. He explores teaching methodologies to explain how to teach AI ethics and human rights in K-12 learning environments. Particular emphasis is made on the survey of existing ethics teaching methodologies and how to adopt existing teaching strategies into AI ethics teaching to improve student understanding about AI ethics and human rights. Demirbilek concludes that danger awaits, with the moral and ethical issues of misuse of AI-powered technologies. There is an increasing demand to establish ethical standards and regulations of AI use. He emphasizes the educational aspect and the need to teach ethical and moral issues which may be caused by AI-powered technologies to future generations. It is important to decorate future generations with AI ethics competencies to critically reflect on their learnings in their future jobs.

In Chapter 6, “Digital Equity,” Patrick Flanagan discusses digital equity through the lens of the digital divide. The digital divide is first contextualized within the coronavirus pandemic to illustrate how digital inequities came to the forefront of people’s agenda during this period. It then moves to discuss the digital divide defining the complex term and offering critical data to illustrate the areas of the world most impacted by this unfortunate reality. Different organizations and groups have made significant moves to narrow the digital gap. These strategies are discussed, however, Flanagan makes the point that none of these groups will be entirely successful if they are not concerned with digital equity. To resolve the threat of the continuance of the virus, vulnerable communities were targeted to ensure that these people already on the fringes did not suffer more than necessary. Government and health officials made strident efforts in the name of equity to resolve the occurrence of the virus in these edges of society. So, too, the proposal in this chapter has been to esteem the value of equity by addressing the digital needs of those on the peripheries of the global village with greater vigilance.

SECTION 3: DIGITAL ETHICS IN HEALTH CARE

The section dedicated to healthcare calls attention to the heightened regulatory burdens we experience and emphasizes the important implications of digital informed consent. Furthermore, one of the chapters is highlighting the ethical nuances related to digital healthcare interactions.

In Chapter 7, “Ethical and Regulatory Challenges in Emerging Health Technologies,” Samia Rizk considers the advances in biotechnology sciences such as genomics, neuroscience, synthetic biology, and nanoscience combined with the rapid developments in computer and data analytics, and the challenges of the increasing complexity of ethical issues associated with modern health technology. Health technology is rapidly evolving and transforming the healthcare system. Data-driven systems are particularly changing traditional healthcare delivery. The gap in knowledge and expertise between these innovative interventions and the existing healthcare contexts for both givers and users creates several ethical, regulatory, and economic challenges which could destabilize trust in the safety, fairness, and effectiveness of the healthcare system. She argues for the need for specialized expertise and wider scope of analysis and a new subfield of ethics termed technology ethics or “techno-ethics” which deals with the framing of principles and methods to guide technology implementation and use.

In Chapter 8, “Ethical Benefits and Drawbacks of Digital Informed Consent,” Wendy Charles considers how the emergence of digital methods has created opportunities to both enhance ethical protections and detract from intended protections. The chapter primarily focuses on informed consent processes that require a complex exchange of information, such as informed consent for healthcare delivery and participation in human subjects research. The authors’ stance throughout this chapter is that the adoption of emerging digital technologies for informed consent does not alter ethical principles, since these retain the moral compass shaped by societal values rooted in autonomy and justice. Digital informed consent methods instead require adaptations of consent processes and appropriate uses of permissioned data to adhere to ethical principles.

Charles makes the case that moving forward, there must not only be laws and regulations about digital informed consent but guidelines and interpretations. The current norms and practices have largely focused on security but require additional guidelines about justifiable processing and terms regarding when re-consent is required to go beyond the initial terms of the agreement. The growing use of digital data requires ongoing education and a shared commitment toward addressing issues and cooperative approaches to protecting ethical protections.

She argues for an individual-centered process that places the well-being and interests of the individual at the forefront and uses the technology to overcome a person's limitations to make a truly informed decision. Organizations designing digital informed consent methods must assume responsibility for obtaining agreement in an ethical manner. They can begin by creating ethical frameworks that delineate an organization's values, detail ethical assessments, and risk mitigation strategies that create an environment more likely to promote a trustworthy agreement and integrity with data protection and use. As digital technologies evolve, laws, regulations, and guidelines must change accordingly to maintain the individual at the center.

In Chapter 9, "Going Telemental: Contact and Intimacy in Digital Mental Health," Shaun Respess looks at telemental health (TMH) which offers unique techniques for persons to interact without needing to be physically present. He examines the long-term advantages and fit alongside conventional therapeutic methods. The author applies elements from care ethics to explore and critique the sustainability of TMH, arguing that such services can compromise the quality of care even while providing several benefits. Furthermore, the author suggests that critical examinations of 'contact', 'intimacy', and embodied spaces should be crucial features of therapeutic assessment, including evaluations of TMH.

He notes that while early indications demonstrate that TMH is effective at reaching more remote populations and could accommodate those who are suddenly displaced or unable to participate in conventional meetings (such as those isolated during a pandemic). Yet, in bridging these gaps the rollout of TMH services must prevent creating new exclusions based on technological literacy or environmental security. One of the benefits of TMH for many providers and clients has been participating from the comfort of their own homes. He explains how TMH offers a unique opportunity to understand and navigate the fundamental features of spaces, media, and relationships in everyday mental health practices. These are underappreciated and vital conditions of care that radically affect its quality and meaning to persons. They may also be used to unpack the peculiar forms of disconnection that threaten the health, safety, and ambitions of patients seeking care, including those stimulated by certain treatment methods. He concludes that whether digital technology will be an enabling or restrictive force moving forward is at present unclear, but it is certainly an intriguing source of controversy and exploration.

SECTION 4: DIGITAL ETHICS AND NEW REALITIES

The chapters included in this last section focus on the novel drivers of change in this digital world and each of the authors cover important domains that influence the design and deployment of digital ethics programs. Social media, investors, and the global business ecosystems are all equally powerful in reshaping, recalibrating, and reconfiguring moral conduct at an individual and societal level.

In Chapter 10, “Social Media, Social Media: How a Mega-Industry Was Built by Systematically Monetizing the Exploitation of Innate Human Emotions and What, if Anything, Can Change This,” Foster Fletcher tackles social media and its systematic exploitation of human emotions, reactions, and biases, blending news and content in one feed to keep users ‘in-app’ and using powerful algorithms to promote more provocative posts, filter content and trigger the reward centers of our brains. He then explores the potential of decentralized technologies, into new social media applications, providing new expectations of user privacy, tighter regulation, and a more equitable monetization system. He speculates that social networks that successfully protect their users’ privacy, reduce hate speech, block bots, reward contributions, and respect personal data have a future. He concludes that people will reject the current social media paradigm and opt for a decentralized, distributed, and equitable future.

In Chapter 11, “Digital Ethics in Technology and Investments,” Ritesh Jain focuses on aspects of Digital Ethics in technology and the growing risks related to data. He provides an overview of the data-related challenges and their ethical uses by organizations and people, as well as emerging technology like AI and its impact on ethical challenges. Based on this he examines the potential ethical challenges of technology that investors should consider and focus on the criticality of the framework requirement and its implementation within businesses to make the right decisions. Finally, Jain lays out the view on ethics and regulations and why companies should commit to ethical practices for their growth. Organizations need to continuously educate ethical decision-making, reinforce, monitor, and empower employees to question any potential unethical issues.

In Chapter 12, “Business Ethics,” Ingrid Vasiliu-Feltes highlights the importance of transforming our conceptualization of business ethics in the digital era and the opportunities related to an optimal design of sustainable digital business ethics programs in this new hyper-connected, hyper-automated digital world. The complex issues of this revised business ethics model are addressed from three perspectives: corporate governance, leadership, and society. The sections related to corporate governance highlight the operational challenges when aiming to incorporate ethics in the board room’s DNA and will emphasize the sustainability imperative ethical business leaders are facing in this digital era. Throughout the chapter, an emphasis is placed on the crucial importance of ethical leadership in this digital era and the unique characteristics required for long-term success.

The author showcases the profound and complex relationship between digital business ethics and other aspects of society, as well as arguing that leaders of the digital era will be expected to embody a hybrid business and digital ethics acumen. This chapter also posits that by adopting a design thinking approach for business ethics in this digital era we can leverage all the benefits offered by emerging technologies and scientific advances while maintaining a human-centric stance. The author concludes that Developing a Universal Code of Conduct for Digital Business Ethics and a Global Digital Ethics Framework should be a moral imperative. As digital ethics advocates we must hold ourselves accountable in order to build a digital ethics legacy for future generations.

In Chapter 13, “Ethical Challenges for Blockchain and Decentralised Finance (DeFi),” Thomason discussed the rapidly growing field of Blockchain and Decentralised Finance. She considers the potential for inadvertent or deliberate automation of unethical conduct at scale and highlights ethical dilemmas for developers, investors, consumers, and regulators at the technology, application, and societal levels. The chapter provides a perspective on the emerging field of DeFi and Blockchain in financial services, a reflection on the ethical questions that arise, how they are being addressed, the key issues, and further research needed in this growing field of Blockchain ethics.

She makes the case for a systematic, cohesive, and joint research agenda informed by stakeholders' views and roles in conceptualizing, developing, and delivering Blockchain technologies, and research at the intersection of financial technologies and their adoption, concerning building capabilities in dealing with externalities such as legal and compliance issues. Thomason argues that new and powerful technologies demand a new approach and one that is co-developed with industry. Further research is needed into the development of and better use of ethical frameworks and criteria to ensure technology is building out in an inclusive, systemic way to address the issues it is supposed to solve.

Even though the potential for Blockchain to transform many aspects of the world is there, there remains a need to ensure that the technology is built and deployed with due concern for ethics. Blockchain can have ethical impacts at the technology, application, and societal levels. It is important that these are considered and built into system design with intentionality. While the promise of automation and decentralization is attractive, it is important to avoid the inadvertent facilitation of unethical conduct. Blockchain technology is a conditional good; it is only as beneficial and useful as the care that is taken to make it.

In Chapter 14, "Ethical Risks in the Cross-Section of Extended Reality (XR), Geographic Information Systems (GIS), and Artificial Intelligence (AI)," Manolova explores the expansion of extended reality, geographic information systems, and artificial intelligence. She makes the point that the mainstreaming of data science has been accompanied by expanded daily hours in front of screens and the emergence of a brand-new internet minute, which contains within itself not only an addicting human-digital symbiosis, but fundamental dependence and emotional investment into digital solutions as a bridge between families, friends, communities, and societies.

She notes that the generational gaps in digital product consumption are constantly growing in sync with emerging technologies as how society consumes information (streaming, social networks, digital realities, immersive environment) and participates in economic exchanges (freelancer, influencer economics) continue to evolve along with cutting edge technologies, which were not available even five years ago. The addictive nature of the internet world transforming the internet minute into the first generation of internet lifetimes, where significant milestones in the human experience could exist almost entirely online. Large demographics already buy in the notions of reality TV, streaming entertainment, social media as forms of escapism and the lines between the real world and the digital one is becoming blurrier with the convergence between social networks and geo-locations.

She argues that regulations do not capture all dimensions of human-machine interaction and data consumption on all sides of the marketplace is fundamental for the establishment of sustainable ethical standards in line with the necessary accountability for the benefit of future generations. The accelerated digitalization of the current data generation is conducive to the rise of extended reality technologies, enhanced machine learning, social network engineering, and the development of new economic constructs, which could have unintended side effects. While these innovations on their own are impressive, their combined effects would once more transform the daily existence of the regular human being.

Manolova makes the point that data science and innovation practitioners within the field of AI, need to maintain integrity in the face of quick and easy business solutions and to treat every user of the emerging technologies as a multi-dimensional human being and not a number on a dashboard, removed from implications outside of the digital domain. The suppliers of the AI should reflect on the side effects of the solutions in the concrete scenario as with XR solutions there are use cases and data streams vastly different from the ones in autonomous vehicles. The coders need to understand the ethical blind spots of the efficiency algorithms that result in their product outputs. The users need to be aware of the invisible data streams that are likely to be captured by a more sophisticated AI algorithm. As further use cases arise, the need for ethical approaches in the design of both environmental and user-interactive application of XR will grow and transform, and there are additional cross sections, which need to be examined considering the constantly evolving designs of technologies and their plethora of applications – there are potential intersections with sociology and economics, and marketing, and psychology. Within these changes the future generations will also change and even the perception of what ethics is and ought to be might transform to reflect the premises of this 6.0 economics of the senses.

In Chapter 15, "My Boss Is an Algorithm: Discussions on How to Best Prepare Students on the Ethics of Human-Machine Interactions at Work," Cynthia Montaudon Tomas explores how employers can address whether or not algorithms will be able to treat people fairly. She reports on a study centered on a private higher education institution in central Mexico. General conditions regarding stress and working hours in the country are described to create the general background of the study, along with two significant regulations that legislate psychosocial risks and remote work. The population considered were full-time faculty members who had moved their activities online. Results were analyzed as a whole and later on divided according to gender to determine whether there were significant differences in terms of burnout syndrome in faculty. She concludes that the digital literacy that is instrumental to educational, working, personal, social lives, and algorithmic literacy needs to include the ethical design of algorithms to make sure that decisions are fair and transparent. One of the critical issues is how to prepare students for a world of machine learning in which algorithms can be used to make work more efficient, and at the same time, ethical, trustworthy, and fair.

In Chapter 16, "Self-Aware Machines," Robin Craig discusses the ethical questions in terms of both how to ensure that self-aware machines will act in the interests of human beings and what rights such machines could have themselves. The chapter describes the general principles of AI ethics; the likelihood of self-aware machines being created; the implications of imposing pro-human ethical constraints upon them; the critical need this creates for an objective ethical system; a proposed objective ethical system, including its implications for the ethical relationship between humans and artificial minds and the further implications this holds for future research in both ethics and computing.

He concludes that while currently, people hold highly diverse views of proper ethics, as artificial intelligence improves it should be possible to encode accurate reasoning ability. That would enable using such machines to investigate the existence and nature of objective ethics, with the potential to not only improve human ethics but also determine to what extent such ethics would regulate the actions of more advanced, self-aware machines. Therefore, such research will be vital for assessing and mitigating the risks to humanity of creating superhuman artificial intelligence.

In Chapter 16, "Ethics, Digital Rights Management, and Cyber Security," Ali Hussain explores Digital Rights Management (DRM) practices for creating better and ethically safe online spaces. He discusses the state-of-the-art DRM paradigms, critically discusses their technical performance, flexibility, and immutability challenges, and critically analyzing the cloud technology stack and ethical features. The integrity of online content lacks some features compared to the real adoption and deployment challenges, such as interoperability, internationalization, and internet policy and argues that research on this issue is needed. He highlights the need to find new and robust and tamper-proof authentication and authorization methods for safer and secure internet. He concludes that the development of an advanced version of the DRM System suitable for cloud computing would enable researchers to protect online content more effectively, with greater content security.

CHALLENGES AND OPPORTUNITIES

There are a plethora of challenges we encounter when deploying digital ethics which range from legal, regulatory, compliance, cyberattacks, cultural, religious, to a financial and economic divide. In our quest to establish global ethical governance, we must foster creativity in deploying novel methods and solutions. One of the options is to apply design thinking methodology for digital ethics.

Before the onset of this latest global pandemic, experts were already predicting an exponential adoption of digital technologies and encouraging businesses across all industries to engage in the digital transformation journey. With the dramatic worldwide economic and workforce changes that we have been experiencing as a consequence of this pandemic, the predicted timeline has markedly accelerated and it has now become imperative for companies to start or accelerate their digital transformation process. Embedding digital ethics into the broader enterprise digital transformation strategy is now considered critical to remain viable and to retain a competitive advantage in a highly disrupted, digital, and virtualized environment. Even during favorable economic times, the road to a successful digital transformation and creating a sustainable culture of digital ethics was filled with hurdles and challenges, requiring a comprehensive strategy and disciplined deployment. Given the financial pressures and disruption in business operations due to the pandemic, it is now even more difficult for companies to accomplish these massive complex tasks. There are several methods to accomplish this, however, the application of design thinking principles seems uniquely suited for a volatile and complex post-pandemic ecosystem due to its agile and human-centered approach. Seven core stages define design thinking methodology and can be used in digital ethics program deployments: empathy, definition, ideation, prototyping, selection, implementation, and feedback. It allows for increased speed of implementation, improved user satisfaction, and cost savings. A design thinking-powered strategy would aim to change the ethics culture by shifting to a human-centered mindset, encourage creative confidence in all employees and the leadership team, express empathy for all stakeholders, and embrace uncertainty. Throughout this process, it would be essential to maintain transparency and open communication with all employees as that will facilitate engagement, acceptance, and successful sustainable deployment.

Leaders that wish to be prepared for the 5th industrial revolution and successful in managing its impact on their organizations and on society will be required to display a complex armamentarium of novel skills, such as technology literacy, technology fluency, and mastery of applied ethics. It used to be sufficient to be a "tech-savvy" leader, however, with the current exponential advancements of a variety of modern technologies it is now imperative for leaders to have a higher degree of technical acumen to earn trust and remain competitive. More importantly, it has become evident that an exponential and abundance mindset will also demand versatility in foundational ethical concepts. Ethics has been an important discipline for centuries, which after decades of marginalization is currently witnessing a resurgence within the scientific and business community due to the complex ethical issues we are facing due to a high degree of automation and augmented intelligence infused in our daily activities. The scientific and business communities, numerous not for profit- and government agencies are all appropriately concerned about ethical issues unique to a digital business ecosystem, with topics like algorithmic bias, social determinants discrimination, data privacy, data ownership, AI transparency and trust making the headlines on a daily basis. Reactive or mitigation approaches are not an optimal solution and one would hope that ethical leadership in this era of the 4th (and soon 5th) industrial revolution will be defined by state of the art strategic planning, complemented by a revision of our education system and a profound transformation of our current employment system.

FUTURE DIRECTIONS

The trend towards automation and digitalization has been accelerated due to the pandemic, with sectors being forced to function virtually. This has stimulated a greater focus on the issue of digital ethics. The World Economic Forum https://www.weforum.org/projects/responsible-use-of-technology has published a report on the responsible use of technology and has cautioned against a "techlash" we might experience due to societal mistrust. Others are also working on these issues including Algorithm Watch and their AI Ethics Global Inventory. Nesta developed a Code of Standards for Public Sector Algorithmic Decision Making.

The Ethics Centre has defined principles to incorporate ethics by design, these are:Ought before can; Net benefit; Non-instrumentalism; Fairness; Self-determination; Accessibility; Responsibility and Purpose. The Blockchain Ethical Design Framework proposes identifying the outcomes and the ethical approach will guide Blockchain design choices. For example, in an aid-distribution Blockchain, the ethical approach may be to ensure that all members of a community have equal access to aid. If the community has significant power disparities among its members, the guiding design philosophy would be to prioritize design choices that minimize disparities in aid distribution. Addressing these questions at the outset of the design process provides ethical intentionality that offers a guiding star to help navigate the inevitable design tradeoffs.

A newly-designed global ethics framework customized for the digital era would include the impact on society, the environment, long term sustainability, education, employment, and emphasis on ESGs. Given the status of the global economy and the impact ethics has on our society, there is an opportunity for exemplary digital ethics leadership. Ethically-driven leaders in the digital era can facilitate the creation and enforce adherence to global digital ethics standards that can positively influence the United Nations SDG agenda for 2030. There needs to be the development of, and better use of ethical frameworks and criteria to ensure technology is building out in an inclusive, systemic way to address the issues supposed to solve.

FUTURE RESEARCH

Other areas for future research including, presenting information via adaptable user interfaces should determine which digital features improve comprehension of the material and perceptions of voluntariness; understanding individuals' ethical concerns and implications of collecting, storing, and sharing individuals' data and additional studies could identify ethical risks and concerns that had not been previously considered for digital informed consent. These results could then inform organizations' ethical frameworks and provide direction for risk mitigation strategies.

Future research will hopefully be centered around the impact of emerging technologies on shaping education, employment and further enhancing ESG-consciousness. Novel technologies and scientific discoveries are emerging daily and challenge us to better understand the influence on societal and organizational behaviors. New technologies such as IoT, BCIs, 6G or 10G, Web 3.0, or the Metaverse will completely disrupt the way we live, the way we work, and the way we entertain ourselves triggering a need for ongoing innovation, transformation, and quality improvement in our global digital ethics ecosystem.

CONCLUSION

As highlighted by various chapters throughout the book, emerging technologies are strong drivers of innovation and transformation, however, they can also act as gateways for unintended negative ethical breaches or as powerful weapons when used by those with malicious intent. As a society, we must focus on the opportunity to seek alignment of digital ethics programs with compliance, regulatory and ESG-conscious policies. This will require a high degree of synchronization and harmonization among all global stakeholders to achieve long-term success.

Several international not-for-profit organizations have outlined the need to establish a global ethical governance to ensure a sustainable digital ethics ecosystem. The strategic roadmap, an industry-agnostic implementation toolkit, as well as collaboration via consortiums or task forces at a global level, will be required to change the paradigm and create a global culture of digital ethics. Some experts have even advocated for the creation of a Universal Code of Digital Ethics to guide legislators, compliance specialists, regulators, and industry leaders in implementing the values outlined by proactive digital ethics programs. We hope that the research and ideas of the contributors to this book, serve to advance thinking and catalyse action, so that ethics becomes foundational to technological design and execution.

Compilation of References

Aboujaoude, E., Salame, W., & Naim, L. (2015). Telemental health: A status update. *World Psychiatry; Official Journal of the World Psychiatric Association (WPA)*, *14*(2), 223–230. doi:10.1002/wps.20218 PMID:26043340

ACM Code 2018 Task Force. (2018). *ACM Code of ethics and professional conduct*. Association for Computing Machinery Committee on Professional Ethics. https://www.acm.org/code-of-ethics

Adams, S. M., Rice, M. J., Jones, S. L., Herzog, E., Mackenzie, L. J., & Oleck, L. G. (2018). TeleMental Health: Standards, Reimbursement, and Interstate Practice. *Journal of the American Psychiatric Nurses Association*, *24*(4), 295–305. doi:10.1177/1078390318763963 PMID:29589800

Adekanmbi, G., & Boitshwarelo, B. (2012). Collaboration in Distance Education in Sub-Saharan Africa: Trends, Trials and. *Cases on Cultural Implications and Considerations in Online Learning*, 375.

Advent Health Medicare Advantage. (2021). *Advent Health Medicare Advantage*. healthgolds.com

Aggarwal, N. (2020, September 19). Introduction to the special issue on Intercultural Digital Ethics. *Philosophy & Technology*. https://link.springer.com/article/10.1007/s13347-020-00428-1

Aggarwal, N. (2021). The norms of algorithmic credit scoring. *The Cambridge Law Journal*, *80*(1), 42–73. doi:10.1017/S0008197321000015

Aguirre, E., Mahr, D., Grewal, D., De Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34–49. doi:10.1016/j.jretai.2014.09.005

AI-Driven platform for content curation — or a Netflix of knowledge! (2017, July 16). *Bold Business*. https://www.boldbusiness.com/human-achievement/ai-driven-platform-content-curation/

Aiken, M. (2016). *The Cyber Effect: A Pioneering Cyberpsychologist Explains how Human Behavior Changes Online*. Random House.

Aisch, G. (2011). *Global digital divide*. Retrieved May 10, 2021, from http://old.driven-by-data.net/about/global-digital-divide/#/0

Ajunwa, I. (2019). The paradox of automation as anti-bias intervention. *Cardozo Law Review*, *41*, 1671–1742. https://scholarship.law.unc.edu/faculty_publications/491

Alampay, E. (2006). Beyond access to ICTs: Measuring capabilities in the information society. *International Journal of Education and Development Using ICT, 2*(3).

Alford, H. J., & Naughton, M. J. (2001). *Managing as if faith mattered: Christian social principles in the modern organization*. University of Notre Dame Press.

Altexsoft. (2019, October 3). *How to successfully implement HR analytics and people analytics in a company*. https://www.altexsoft.com/blog/how-to-implement-hr-analytics/

AMA Principles of Medical Ethics. (1995). *American Medical Association*. https://www.ama-assn.org/about/publications-newsletters/ama-principles-medical-ethics

American Medical Association. (2016). *Code of medical ethics overview*. https://www.ama-assn.org/delivering-care/ethics/code-medical-ethics-overview

American Psychological Association. (2020, June 5). *Psychologists Embrace Telehealth to Prevent the Spread of COVID-19*. https://www.apaservices.org/practice/legal/technology/psychologists-embrace-telehealth

Anabo, I. F., Elexpuru-Albizuri, I., & Villardón-Gallego, L. (2019). Revisiting the Belmont Report's ethical principles in internet-mediated research: Perspectives from disciplinary associations in the social sciences. *Ethics and Information Technology*, *21*(2), 137–149. doi:10.100710676-018-9495-z

Anderson, D., Bonaguro, J., McKinney, M., Nicklin, A., & Wiseman, J. (2018). *Ethics & algorithms toolkit*. http://ethicstoolkit.ai/

Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, *28*(4), 15–26. doi:10.1609/aimag.v28i4.2065

Anderson, M., & Anderson, S. L. (2015). Toward ensuring ethical behavior from autonomous systems: A case-supported principle-based paradigm. *Industrial Robot: An International Journal*, *42*(4), 324–331. doi:10.1108/IR-12-2014-0434

Anderson, M., Anderson, S. L., & Armen, C. (2004). Towards machine ethics. In *Proceedings of the AAAI-04 Workshop on Agent Publishers: Theory and Practice* (pp. 2-7). AAAI Press.

Antonolpoulos, F., Petrakis, E. G. M., Sotiriadis, S., & Bessis, N. (2018). A physical access control system on the cloud. *Procedia Computer Science*, *130*, 318–325. doi:10.1016/j.procs.2018.04.045

Anyoha, R. (2017, August 28). The history of artificial intelligence. *Science in the News*. https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

Arenas-Torres, F., Bustamante-Ubilla, M., & Campos-Troncoso, R. (2021). *The incidence of social responsibility in the adoption of business practices*. MDPI. https://www.mdpi.com/2071-1050/13/5/2794

Arisona, S., Mueller, P., Meriaux, A., & Hansen, R. (2020, March 10-13). *Extended reality (XR) with ArcGIS* [Paper presentation]. 2020 Esri Developer Summit, Palm Springs, CA, United States. https://www.esri.com/content/dam/esrisites/en-us/events/conferences/2020/developer-summit/extended-reality-xr-with-arcgis.pdf

Arkin, R. C., Ulam, P., & Wagner, A. R. (2011). Moral decision making in autonomous systems: Enforcement, moral emotions, dignity, trust, and deception. *Proceedings of the IEEE*, *100*(3), 571–589. doi:10.1109/JPROC.2011.2173265

Arocho Esteves, J. (2020, February 28). *Humanity, ethics must be at the center of AI technology, pope says*. Crux. https://cruxnow.com/vatican/2020/02/humanity-ethics-must-be-at-center-of-ai-technology-pope-says/

Artificial intelligence. (2021, August 3). In *Wikipedia*. https://en.wikipedia.org/wiki/Artificial_intelligence

Artyushina, A. (2020). *Is civic data governance the key to Democratic Smart Cities? the role of the Urban Data Trust in sidewalk toronto*. Telematics and Informatics. https://www.sciencedirect.com/science/article/pii/S0736585320301155

Association for Computing Machinery. (2018, June 22). *ACM code of ethics and professional conduct*. https://www.acm.org/code-of-ethics

Athey, S., Catalini, C., & Tucker, C. E. (2017, September 27). *The digital privacy paradox: Small money, small costs, small talk*. National Bureau of Economic Research. https://www.google.com/books/edition/The_Digital_Privacy_Paradox/2AFfswEACAAJ?hl=en

Baier, C. (1970). *Challenges in the implementation of responsible business conduct*. OPUS 4 | Challenges in the implementation of responsible business conduct. https://opus4.kobv.de/opus4-ku-eichstaett/frontdoor/index/index/docId/674

Baier, A. (1986). Trust and Antitrust. *Ethics*, *96*(2), 231–260. doi:10.1086/292745

Bailenson, J. (2018). Protecting nonverbal data tracked in virtual reality. *JAMA Pediatrics*, *172*(10), 905–906. doi:10.1001/jamapediatrics.2018.1909 PMID:30083770

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., & Slater, M. (2020). Virtual body ownership and its consequences for implicit racial bias are dependent on social context. *Royal Society Open Science*, *7*(12), 201848. doi:10.1098/rsos.201848 PMID:33489296

Banta, D., Kristensen, F. B., & Jonsson, E. (2009). A history of health technology assessment at the European level. *International Journal of Technology Assessment in Health Care*, *25*(S1), 68–73. Advance online publication. doi:10.1017/S0266462309090448 PMID:19534837

Bardi, J. (2018, November 11). *3 secrets to creating immersive virtual environments with Unity and Vuforia*. Marxent Labs. https://www.marxentlabs.com/virtual-environments-unity/

Barker, P. (2011). *Mental Health Ethics: The human context*. Routledge.

Barnes, M. (2012). *Care in Everyday Life: An Ethic of Care in Practice*. Policy Press.

Barratt, T., Veen, A., & Goods, C. (2020, August 24). *Algorithms workers can't see are increasingly pulling the management strings*. The Conversation. https://theconversation.com/algorithms-workers-cant-see-are-increasingly-pulling-the-management-strings-144724

Barrera, A. Z., Dunn, L. B., Nichols, A., Reardon, S., & Muñoz, R. F. (2016). Getting it "right:" Ensuring informed consent for an online clinical trial. *Journal of Empirical Research on Human Research Ethics; JERHRE*, *11*(4), 291–298. doi:10.1177/1556264616668974 PMID:27630213

Barrett, C., Dill, D. L., Kochenderfer, M. J., & Sadigh, D. (n.d.). *Stanford Center for AI Safety. White Paper.* Available in http://aisafety.stanford.edu/

Barzilai-Nahon, K. (2006). Gaps and bits: Conceptualizing measurements for digital divide/s. *The Information Society*, *22*(5), 269–278. doi:10.1080/01972240600903953

Bashshur, R. L., Howell, J. D., Krupinski, E. A., Harms, K. M., Bashshur, N., & Doarn, C. R. (2016). The Empirical Foundations of Telemedicine Interventions in Primary Care. *Telemedicine Journal and e-Health*, *22*(5), 342–375. doi:10.1089/tmj.2016.0045 PMID:27128779

Batterywala, J., & Agarwal, P. (2018). *Opening the black box: Managing algorithm risks*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/in/Documents/risk/in-risk-Managing-algorithmic-risks-17aug2020-noexp.pdf

BBC News. (2013). *Kenya IT Hubs Launched for Primary Schools*. Retrieved May 10, 2021, from https://www.bbc.co.uk/news/world-africa-24147333

BBC. (2021). *AI at work: Staff 'hired and fired by algorithm'*. https://www.bbc.com/news/technology-56515827

Beard, C. (2020, October 1). *Peace of Mind: There's an App for That.* McLean Hospital. https://www.mcleanhospital.org/essential/peace-mind-theres-app

Beauchamp, T. L., & Childress, J. F. (2013). *Principles of Biomedical Ethics* (8th ed.). Oxford University Press.

Beck, J., & Asher, M. (2021, February 3). Why Decentralized Autonomous Organizations (DAOs) are essential to DeFi. *Consensys*. https://consensys.net/blog/codefi/daos/

Bedi, R. K., Singh, J., & Gupta, S. K. (2018). MWC: An efficient and secure multi-cloud storage approach to leverage augmentation of multi-cloud storage services on mobile devices using fog computing. *The Journal of Supercomputing*. Advance online publication. doi:10.100711227-018-2304-y

Bell, P. (2001). Algorithms: The intellectual capital of the internet. *Ivey Business Journal*. https://iveybusinessjournal.com/publication/algorithms-the-intellectual-capital-of-the-internet/

Benbouzid, B. (2018). Quand prédire, c'est gérer: La police prédictive aux États-Unis [When to predict is to manage: Predictive policing in the United States]. *Reseaux (London)*, *5*(211), 221–256. doi:10.3917/res.211.0221

Bennemann, S. (2016). *Disagreement in ethics*. Freiburg Institute for Advanced Studies. https://www.frias.uni-freiburg.de/en/events/conferences/disagreement-in-ethics

Bergquist, M., Ljungberg, J., & Rolandsson, B. (2011). A Historical Account of the Value of Free and Open Source Software: From Software Commune to Commercial Commons. In Open Source Systems: Grounding Research (pp. 196-207). Springer Berlin Heidelberg.

Bernal Bernabe, J., Marin Perez, J. M., Alcaraz Calero, J. M., Garcia Clemente, F. J., Martinez Perez, G., & Gomez Skarmeta, A. F. (2014). Semantic-aware multi-tenancy authorization system for cloud architectures. *Future Generation Computer Systems*, *32*, 154–167. doi:10.1016/j.future.2012.05.011

Bernstein, A. (2005). *The capitalist manifesto*. University Press of America.

Berry, W. (1971). *The unforeseen wilderness: An essay on Kentucky's Red River gorge*. Counterpoint.

Bhatta, N. (1970). Emerging ethical challenges of leadership in the digital era: A Multi-Vocal literature review. *Electronic Journal of Business Ethics and Organization Studies*. https://jyx.jyu.fi/handle/123456789/74932

Bigham, T., Nair, S., Soral, S., Tua, A., Gallo, V., Lee, M., Mews, T., & Fouché, M. (2018). *AI and risk management: Innovating with confidence*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/financial-services/deloitte-uk-ai-and-risk-management.pdf

Bindra, C. (2021, January 25). *Building a privacy-first future for web advertising*. Google Ads&Commerce Blog. https://blog.google/Products/Ads-Commerce/2021-01-Privacy-Sandbox/

Bioy, H. (2020, June 16). *Do sustainable funds beat their rivals?* Morningstar. https://www.morningstar.co.uk/uk/news/203214/do-sustainable-funds-beat-their-rivals.aspx

Birgisson, A., Gibbs Politz, J., Erlingsson, Ú., Taly, A., Vrable, M., & Lentczner, M. (2014). *Macaroons: Cookies with Contextual Caveats for Decentralized Authorization in the Cloud*. doi:10.14722/ndss.2014.23212

Birrell, E., & Schneider, F. B. (2013). Federated Identity Management Systems: A Privacy-Based Characterization. *IEEE Security and Privacy*, *11*(5), 36–48. doi:10.1109/MSP.2013.114

Black, D. (1979). The paradox of medical care. *JR Co1 Physicians Land, 13*, 57-65.

Blanchfield, L. (2011). The United States Convention on the Rights of the Child: Background and Policy Issues. In K. McGowan (Ed.), United Nations: International Agreements and Efforts. Academic Press.

Blandford, A., Wesson, J., Amalberti, R., AlHazme, R., & Allwihan, R. (2020). Opportunities and challenges for telehealth within, and beyond, a pandemic. *The Lancet. Global Health*, *8*(11), e1364–e1365. doi:10.1016/S2214-109X(20)30362-4 PMID:32791119

Bleier, A., & Eisenbeiss, M. (2015). The importance of trust for personalized online advertising. *Journal of Retailing*, *91*(3), 390–409. doi:10.1016/j.jretai.2015.04.001

Boczkowski, P., Mitchelstein, E., & Matassi, M. (2017). Incidental news: How young people consume news on social media. In T. X. Bui, & R. Sprague Jr. (Eds.), *Proceedings of the 50th Hawaii International Conference On System Sciences* (pp. 1785-1792). Computer Society Press. 10.24251/HICSS.2017.217

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, *103*(1–2), 347–356. doi:10.1016/S0004-3702(98)00055-1

Bogen & Rieke. (2018, Dec.). Help Wanted. An Examination of Hiring Algorithms, Equality and Bias. *Upturn*, 36-38.

Bostrom, N. (2003). Ethical issues in advanced artificial intelligence. In I. Smit, W. Wallach & G. E. Lasker (Eds.), Cognitive, emotive and ethical aspects of decision making in humans and in artificial intelligence (Vol. 2, pp. 12-17). International Institute of Advanced Studies in Systems Research and Cybernetics.

Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *The Cambridge Handbook of Artificial Intelligence* (pp. 316–334). Cambridge University Press. doi:10.1017/CBO9781139046855.020

Bourgais, A., & Ibnouhsein, I. (2021). *Ethics-by-design: The next frontier of industrialization*. AI and Ethics. https://link.springer.com/article/10.1007/s43681-021-00057-0

Bradshaw, S., & Howard, P. N. (2018). *Challenging truth and trust: A global inventory of organized social media manipulation*. University of Oxford. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf

Brall, C., Schröder-Bäck, P., & Maeckelberghe, E. (2019). Ethical aspects of digital health from a justice point of view. *European Journal of Public Health*, *29*(Supplement_3), 18–22. doi:10.1093/eurpub/ckz167 PMID:31738439

Brantingham, P. J., Valasik, M., & Mohler, G. O. (2018). Does predictive policing lead to biased arrests? Results from a randomized controlled trial. *Statistics and Public Policy (Philadelphia, Pa.)*, *5*(1), 1–6. doi:10.1080/2330443X.2018.1438940

Brave Launches Next-Generation Browser that Puts Users in Charge of Their Internet Experience with Unmatched Privacy and Rewards. (2019). https://www.prnewswire.com/news-releases/brave-launches-next-generation-browser-that-puts-users-in-charge-of-their-internet-experience-with-unmatched-privacy-and-rewards-300957360.html

Brayne, S. (2018). The criminal law and law enforcement implications of big data. *Annual Review of Law and Social Science*, *14*(1), 293–308. doi:10.1146/annurev-lawsocsci-101317-030839

Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H., Bähler, J., & Wood, V. (2008). The BioGRID Interaction Database: 2008 update. *Nucleic Acids Research, 36*(1), D637–D640, doi:10.1093/nar/gkm1001

Brenkert, G. G. (2016). Business ethics and Human Rights: An Overview. *Business and human Rights Journal*. https://www.cambridge.org/core/journals/business-and-human-rights-journal/article/business-ethics-and-human-rights-an-overview/4E12322863D6BA2B17871B03EDA9BBB9

Briône, P. (2020, March 6). *My boss the algorithm: An Ethical look at algorithms in the workplace*. Acas. https://www.acas.org.uk/my-boss-the-algorithm-an-ethical-look-at-algorithms-in-the-workplace

Broussard, M. (2019). *Artificial unintelligence: How computers misunderstand the world*.

Brown, T., & Wyatt, J. (2010.). *Design thinking for social innovation*. Development Outreach. https://elibrary.worldbank.org/doi/abs/10.1596/1020-797x_12_1_29

Bucher, E. L., Schou, P. K., & Waldkirch, M. (2021). Pacifying the algorithm–Anticipatory compliance in the face of algorithmic management in the gig economy. *Organization*, *28*(1), 44–67. doi:10.1177/1350508420961531

Burgess, M. (2020, August 20). *The lessons we all must learn from the A-levels algorithm debacle*. Wired. https://www.wired.co.uk/article/gcse-results-alevels-algorithm-explained

Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T. (2017). Ethical considerations in artificial intelligence courses. *AI Magazine*, *38*(2), 22–34. doi:10.1609/aimag.v38i2.2731

Burwell, S., Sample, M., & Racine, E. (2017). *Ethical aspects of brain computer interfaces: A scoping review*. BMC Medical Ethics. https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-017-0220-y

Businesswire. (2018, June 5). *$699 million social robot market - Global forecasts from 2018 to 2023*. https://www.businesswire.com/news/home/20180605006538/en/699-Million-Social-Robot-Market---Global-Forecasts-from-2018-to-2023---ResearchAndMarkets.com

Butler, J. V., Giuliano, P., & Guiso, L. (2015). Trust, values, and false consensus. *International Economic Review*, *56*(3), 889–915. doi:10.1111/iere.12125

California Consumer Privacy Act of 2018. (2018). *Title 1.81., Sections 1798.100 - 1798.199.100*. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

Calo, R., & Citron, D. K. (2021). The automated administrative state: A crisis of legitimacy. *Emory Law Journal*, *70*(4), 797–845. https://scholarlycommons.law.emory.edu/elj/vol70/iss4/1

Capelli, P. (2018, February 20). Are algorithms good managers? *Human Resources Executives*. https://hrexecutive.com/are-algorithms-good-managers/

Captain, S. (2020, July 1). *In 20 years, your boss may track your every glance, keystroke, and heartbeat*. Fast Company. https://www.fastcompany.com/90450122/in-20-years-your-boss-may-track-your-every-glance-keystroke-and-heartbeat

Carey, D., & Smith, M. (2016, April 22). How companies are using simulations, competitions, and analytics to hire. *Harvard Business Review*. https://hbr.org/2016/04/how-companies-are-using-simulations-competitions-and-analytics-to-hire

Carlson-Wee, O. (2017, January 8). *The future is a decentralized internet*. TechCrunch. https://techcrunch.com/2017/01/08/The-Future-Is-A-Decentralized-Internet/

Carpentier, C. L., & Braun, H. (2020). Agenda 2030 for sustainable development: A powerful global framework. *Journal of the International Council for Small Business*, *1*(1), 14–23. doi:10.1080/26437015.2020.1714356

Carrin, G., Evans, D., & Xu, K. (2007, September). Designing health financing policy towards universal coverage. *Bulletin of the World Health Organization*, *85*(9), 649–732. doi:10.2471/BLT.07.046664 PMID:18026615

CBC News. (2011). *Internet 'kill switch' easy target in Egypt*. CBC News. https://www.cbc.ca/news/technology/internet-kill-switch-easy-target-in-egypt-1.1110730

Chadwick, D. W., & Fatema, K. (2012). A privacy preserving authorisation system for the cloud. *Journal of Computer and System Sciences*, *78*(5), 1359–1373. doi:10.1016/j.jcss.2011.12.019

Chadwick, R. (1998). Professional Ethics. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*. Routledge.

Chalil Madathil, K., Koikkara, R., Obeid, J., Greenstein, J. S., Sanderson, I. C., Fryar, K., Moskowitz, J., & Gramopadhye, A. K. (2013). An investigation of the efficacy of electronic consenting interfaces of research permissions management system in a hospital setting. *International Journal of Medical Informatics*, *82*(9), 854–863. doi:10.1016/j.ijmedinf.2013.04.008 PMID:23757370

Chang, V. (2021). *An ethical framework for big data and smart cities*. Technological Forecasting and Social Change. https://www.sciencedirect.com/science/article/abs/pii/S0040162520313858?dgcid=rss_sd_all

Charles Schwab. (n.d.). *Schwab intelligent portfolios*. https://intelligent.schwab.com/

Chen, D. (2021, February 24). *What are blockchain domain NFTs? A full introduction*. https://unstoppabledomains.com/blog/what-are-blockchain-domain-nfts-a-full-introduction

Chen, J. Q. (2019). Who should be the boss? Machines or a human? In P. Griffiths & M. Nowshade Kabir (Eds.), *ECIAIR 2019 European Conference on the Impact of Artificial Intelligence and Robotics* (pp. 71–79). Academic Conferences and Publishing International Limited.

Cheng, Y. Y., Li, H., & Zhang, N. (2016). Character-Based Online Key Management in Cloud Computing Environment. *Proceedings of 2016 Ieee Advanced Information Management, Communicates, Electronic and Automation Control Conference (Imcec 2016)*, 738–741. 10.1109/IMCEC.2016.7867307

Chhin, V., Roussos, J., Michaelson, T., Bana, M., Bezjak, A., Foxcroft, S., Hamilton, J. L., & Liu, F.-F. (2017). Leveraging mobile technology to improve efficiency of the consent-to-treatment process. *JCO Clinical Cancer Informatics*, *1*(1), 1–8. doi:10.1200/CCI.17.00041 PMID:30657388

Chiauzzi, E., & Wicks, P. (2019). Digital trespass: Ethical and terms-of-use violations by researchers accessing data from an online patient community. *Journal of Medical Internet Research*, *21*(2), e11985. doi:10.2196/11985 PMID:30789346

Children's Commissioner's Growing Up Digital Taskforce. (2017). *Growing Up Digital: A Report of the Growing Up Digital Task Force.* Children's Commissioner. https://www.childrenscommissioner.gov.uk/report/growing-up-digital/

Chokngamwong, R., & Jirabutr, N. (2015). Mobile Digital Right Management with Enhanced Security using Limited-Use Session Keys. *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Ecti-Con)*.

Christensen, L. R., Marcik, W., Rafert, G., & Wong, C. (2016). *The global economic impacts associated with virtual and augmented reality*. Analysis Group. https://www.analysisgroup.com/globalassets/content/insights/publishing/analysis_group_vr_economic_impact_report.pdf

Clark, K., Duckham, M., Guillemin, M., Hunter, A., McVernon, J., O'Keefe, C., Pitkin, C., Prawer, S., Sinnott, R., Warr, D., & Waycott, J. (2019). Advancing the ethical use of digital data in human research: Challenges and strategies to promote ethical practice. *Ethics and Information Technology*, *21*(1), 59–73. doi:10.100710676-018-9490-4

Coeckelbergh, M. (2020). *AI Ethics. MIT Press*.

Coeckelbergh, M., & Stahl, B. C. (2016). Ethics of Healthcare Robotics: Towards Responsible Research and Innovation. *Robotics and Autonomous Systems*, *86*, 152–161. doi:10.1016/j.robot.2016.08.018

Cofield, C. (2015, July 21). *Stephen Hawking: Intelligent aliens could destroy humanity, but let's search anyway*. Space. https://www.space.com/29999-stephen-hawking-intelligent-alien-life-danger.html

Cohan, J. A. (2010). Honor killings and the cultural defense. *California Western International Law Journal*, *40*(2), 177–252.

Colle, R. D., & Roman, R. (2001). Editorial: The Telecenter Environment in 2002. *The Journal of Development Communication*, *12*(2), 15.

Çöltekin, A., Griffin, A. L., Slingsby, A., Robinson, A. C., Christophe, S., Rautenbach, V., Chen, M., Pettit, C., & Klippel, A. (2020). Geospatial information visualization and extended reality displays. In H. Guo, M. F. Goodchild, & A. Annoni (Eds.), *Manual of digital earth* (pp. 229–277). Springer. doi:10.1007/978-981-32-9915-3_7

Committee for the Coordination of Statistical Activities. (2020). *How COVID-19 is changing the world: A statistical perspective* (Vol. 1). https://data.unicef.org/wp-content/uploads/2020/07/covid19-report-ccsa.pdf

Compaine, B. M. (2001). *The digital divide: Facing a crisis or creating a myth?* Mit Press. doi:10.7551/mitpress/2419.001.0001

Conrad, P. (2007). *The Medicalization of Society: On the Transformation of Human Conditions into Treatable Disorders*. JHU Press.

Convention on the Rights of the Child. Nov. 20, 1989, 1577 U.N.T.S. 3. www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx

Cook, A., Robinson, M., Ferrag, M. A., Maglaras, L., He, Y., Jones, K., & Janicke, H. (2017). *Internet of Cloud: Security and Privacy issues*. Academic Press.

Coravos, A., Doerr, M., Goldsack, J., & Wood, W. A. (2020). Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. *NPJ Digital Medicine, 3*(1). https://www.researchgate.net/deref/http%3A%2F%2Fwww.nature.com%2Fnpjdigitalmed

Corple, D. J., Zoltowski, C. B., Kenny Feister, M., & Buzzanell, P. M. (2020). Understanding ethical decision-making in design. *Journal of Engineering Education*, *109*(2), 262–280. doi:10.1002/jee.20312

Cossitt, A. (2020). *Why Health Care Organizations Need Technology Ethics Committees. February 5, in Ethics*. Hastings Bioethics Forum, Health, And Health Care. https://www.thehastingscenter.org/

Costello, E. (2017, August 31). *Big banks back Blockchain for back-office business*. International Investment. https://www.internationalinvestment.net/internationalinvestment/news/3504643/banks-blockchain-office-business

Cotter, K. M. (2020). *Critical algorithmic literacy: Power, epistemology, and platforms* (Publication No. 28029214) [Doctoral dissertation, Michigan State University]. ProQuest Dissertation and Thesis Global.

Council of Europe. (2020). *History of artificial intelligence.* https://www.coe.int/en/web/artificial-intelligence/history-of-ai

Cousaert, S. (2021, March 25). *Generalizing knowledge on DEXs with AMMs — Part I.* Medium. https://medium.com/uclcbt/generalizing-knowledge-on-dexs-with-amms-2963d07ebac7

Craig, R. (2010). Good without God. In W. Bonett (Ed.), *The Australian book of atheism* (pp. 349–361). Scribe Publications.

Cuende, L. (2020, June 11). *DAOs, the next big thing after social media.* Aragon. https://aragon.org/blog/daos-the-next-big-thing

Cui, J., Zhou, H., Zhong, H., & Xu, Y. (2018). AKSER: Attribute-based keyword search with efficient revocation in cloud computing. *Information Sciences*, *423*, 343–352. doi:10.1016/j.ins.2017.09.029

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2020). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*, *65*(3), 644–676. doi:10.1177/0001839219867024

Cvetkovich, A. (2012). *Depression: A Public Feeling*. Duke University Press.

Das, A. K., Mishra, D., & Mukhopadhyay, S. (2015). An anonymous and secure biometric-based enterprise digital rights management system for mobile environment. *Security and Communication Networks*, *8*(18), 3383–3404. doi:10.1002ec.1266

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. Available at https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Davis, M. (1993). Ethics across the curriculum: Teaching professional responsibility in technical courses. *Teaching Philosophy*, *16*(3), 205–235. doi:10.5840/teachphil199316344

Davis, M. (2006). Integrating ethics into technical courses: Micro-insertion. *Science and Engineering Ethics*, *12*(4), 717–730. doi:10.100711948-006-0066-z PMID:17199146

De Angelis, F. L., & Di Marzo Serugendo, G. (2017). SmartContent—Self-Protected Context-Aware Active Documents for Mobile Environments. *Electronics (Basel)*, *6*(1), 17. Advance online publication. doi:10.3390/electronics6010017

De Stefano, V. (2020). Algorithmic bosses and what to do about them: Automation, artificial intelligence, and labour protection. In D. Marino & M. Monaca (Eds.), *Economic and policy implications of artificial intelligence* (Vol. 288, pp. 65–86). Springer., doi:10.1007/978-3-030-45340-4_7

De Sutter, E., Zaçe, D., Boccia, S., Di Pietro, M. L., Geerts, D., Borry, P., & Huys, I. (2020). Implementation of electronic informed consent in biomedical research and stakeholders' perspectives: Systematic review. *Journal of Medical Internet Research*, *22*(10), e19129. doi:10.2196/19129 PMID:33030440

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, *358*(6362), 486–492. doi:10.1126cience.aan8871 PMID:29074769

Dehmer, M., & Schafer, P. (2021, August 30). *Workplace analytics insights*. Microsoft. https://docs.microsoft.com/en-us/workplace-analytics/use/insights

Demeke, H. B., Merali, S., Marks, S., Pao, L. Z., Romero, L., Sandhu, P., Clark, H., Clara, A., McDow, K. B., Tindall, E., Campbell, S., Bolton, J., Le, X., Skapik, J. L., Nwaise, I., Rose, M. A., Strona, F. V., Nelson, C., & Siza, C. (2021). Trends in use of telehealth among health centers during the COVID-19 pandemic— United States, June 26–November 6, 2020. *MMWR. Morbidity and Mortality Weekly Report*, *70*(7), 240–244. doi:10.15585/mmwr.mm7007a3 PMID:33600385

Deng, Z., Li, K., Li, K., & Zhou, J. (2017). A multi-user searchable encryption scheme with keyword authorization in a cloud storage. *Future Generation Computer Systems*, *72*, 208–218. doi:10.1016/j.future.2016.05.017

Deweerdt, S. (2019). Deep connections. *Nature*, *571*(7766), S6–S8. doi:10.1038/d41586-019-02208-0 PMID:31341309

Dickson, B. (2017, October 8). *Can blockchain decentralize the internet*. VentureBeat. https://venturebeat.com/2017/10/08/Can-Blockchain-Decentralize-The-Internet/

Dickson, B. (2019, May 20). *Artificial intelligence created filter bubbles. Now it's helping to fight it*. TechTalks. https://bdtechtalks.com/2019/05/20/artificial-intelligence-filter-bubbles-news-bias/

Dierksmeier, C., & Seele, P. (2016). Cryptocurrencies and business ethics. *Journal of Business Ethics*, *152*(1), 1–14. doi:10.100710551-016-3298-0 PMID:30930508

Digital-data studies need consent. (2019). *Nature, 572*(7767), 5–5.

Dike, C. C., Candilis, P., Kocsis, B., Sidhu, N., & Recupero, P. (2019). Ethical considerations regarding internet searches for patient information. *Psychiatric Services (Washington, D.C.)*, *70*(4), 324–328. doi:10.1176/appi.ps.201800495 PMID:30651058

DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality. *Social Inequality: From Unequal Access to Differentiated Use*, 355-400.

Donkor, C., Slobodjanjuk, A., Cremer, K., & Weisshaar, J. (2017). *The way we work–in 2025 and beyond*. PwC. https://www.pwc.ch/en/publications/2017/the-way-we-work-hr-today_pwc-en_2017.pdf

Dotan, R. (2020). Theory choice, non-epistemic values, and machine learning. *Synthese*. Advance online publication. doi:10.100711229-020-02773-2

Draper, B. L., & Owen, S. (2016). *Big Data Research: Practical Solutions to Emerging Challenges for IRBs.* Webinar of the Public Responsibility in Medicine and Research (PRIM&R) organization. www.PRIM&R.org

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2019). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal*, *30*(1), 114–132. doi:10.1111/1748-8583.12258

Dzieza, J. (2020, February 27). *How hard will the robots make us work?* The Verge. https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon

Economic and Social Council. (2017). *Building digital competencies to benefit from existing and emerging technologies with a special focus on gender and youth dimensions*. United Nations. https://unctad.org/system/files/official-document/ecn162018d3_en.pdf

Edenberg, E., & Jones, M. L. (2019). Analyzing the legal roots and moral core of digital consent. *New Media & Society*, *21*(8), 1804–1823. doi:10.1177/1461444819831321

Edis, T. (1998). How Gödel's theorem supports the possibility of machine intelligence. *Minds and Machines*, *8*(2), 251–262. doi:10.1023/A:1008233720449

Eggers, W.D., Torley, M., & Kishnani, P. (2018). *Principles of regulating emerging technologies*. A report from the Deloitte center for government insights. the future of regulation. https://www2.deloitte.com/us/en/insights.html

Elafris. (2020). *People get tired and bored and make mistakes, but algorithms don't*. https://www.elafris.com/blog-post/people-get-tired-and-bored-and-make-mistakes-but-algorithms-dont/

Electronic Signatures in Global and National Commerce Act. (1999). *15 U.S.C. Ch. 96, Sections 7001-7031 of Public Law 106–229*. https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter96&edition=prelim

Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, *5*, 40–60. doi:10.17351/ests2019.260

Engster, D. (2007). *The Heart of Justice: Care Ethics and Political Theory*. Oxford University Press.

Esposito, C. (2018). Interoperable, dynamic and privacy-preserving access control for cloud data storage when integrating heterogeneous organizations. *Journal of Network and Computer Applications*, *108*, 124–136. doi:10.1016/j.jnca.2018.01.017

Estes, A. (2011). The U.N. declares Internet access a human right. *National Journal.* Retrieved May 10, 2021, from https://www.nationaljournal.com/dailyfray/the-u-n-declares-internet-access-a-human-right-20110606

Ethics. (2021). *Oxford English Dictionary*. Retrieved from http://oxforddictionaries.com/definition/english/ethics

EU Agency for Fundamental Rights. (2020). *Getting the Future Right*. Artificial Intelligence and Fundamental Rights.

EUnetHTA Project. (2008). *Encyclopedia of Public Health* (W. Kirch, Ed.). Springer. doi:10.1007/978-1-4020-5614-7_1066

EUnetHTA. (2013). *The HTA Core Model is a methodological framework for shared production and sharing of HTA information*. The HTA Core Model is a registered trademark. Joint Action 2, Work Package 8. HTA Core Model ® version 2.0. HTTP://WWW.COREHTA.INFO/BROWSEMODEL.ASPX

EUnetHTA. (n.d.). *Creating, Facilitating, and Promoting Sustainable Health Technology Assessment (HTA) Cooperation in Europe News_FF9900-500*. Retrieved March 29, 2021, from https://www.eunethta.eu

EUR-Lex. (2018). *The general data protection regulation applies in all Member States from 25 May 2018*. https://eur-lex.europa.eu/

European Commission. (2018). *Data protection in the EU*. https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

European Commission. (n.d.). *Data protection: Rules for the protection of personal data inside and outside the EU*. https://ec.europa.eu/info/law/law-topic/data-protection_en

European Council of Medical Orders. (2011, June 10). *European charter of medical ethics*. http://www.ceom-ecmo.eu/en/view/principles-of-european-medical-ethics

European Group on Ethics in Science and New Technologies (EGE). (2021). https://ec.europa.eu/info/research-and-innovation/strategy/support-policy-making/scientific-support-eu-policies/

European Medicines Agency (EMA). (2020). *Qualification of digital technology-based methodologies to support approval of medicinal products*. EMA/219860/2020. https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance/qualification-novelmethodologies-

European Parliament and Council of the European Union. (2016a, April 14). *GDPR, Article 4 - Definitions*. General Data Protection Regulation 2016/679. https://gdpr.eu/article-4-definitions/

European Parliament and Council of the European Union. (2016b, April 14). *GDPR, Article 7 - Conditions for consent*. General Data Protection Regulation 2016/679. https://gdpr.eu/article-7-how-to-get-consent-to-collect-personal-data/

European Parliament and Council of the European Union. (2016c, April 14). *GDPR, Recital 32 - Conditions for consent*. General Data Protection Regulation 2016/679. https://gdpr.eu/recital-32-conditions-for-consent/

Evans, C. (2012). Five reasons we have yet to close the digital divide. *Microsoft in Education Blog*. Retrieved May 10, 2021, from, http://blogs.technet.com/b/microsoft_in_education/archive/2012/02/02/five-reasons-we-have-yet-to-close-the-digital-divide.aspx

Facebook to acquire Oculus. (2014). https://about.fb.com/news/2014/03/facebook-to-acquire-oculus/

Fagan, G. (2020). Roman violence: Attitudes and practice. In G. Fagan, L. Fibiger, M. Hudson, & M. Trundle (Eds.), *The Cambridge world history of violence* (pp. 550–571). Cambridge University Press. doi:10.1017/9781316341247.029

Fanning, P. (2020, April 8). Connected working: Human-machine interaction. *Eureka!* https://www.eurekamagazine.co.uk/design-engineering-features/technology/connected-working-human-machine-interaction/225975/

Faraj, S. (2019, September 15). What the future of work holds in the age of the learning algorithm. *Delve*. https://www.mcgill.ca/delve/article/blog/what-future-work-holds-age-learning-algorithm

Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, *16*(8). Advance online publication. doi:10.1111/phc3.12760

FDA. (1976). Medical Device Amendments of 1976. In *History of Federal Regulation: 1902–Present. Major legislation with regard to drugs and medical devices*. www.fdareview.org/

FDA. (2018). *Software as a medical device. An agile model for food and drug administration (FDA)-regulated software in health care*. https://www2.deloitte.com/

FDA. (2019). *De Novo Classification Request 11/20*. https://www.fda.gov/medical-devices

FDA. (2019). *Premarket Approval (PMA). 05/16*. https://www.fda.gov/medical-devices/

FDA. (2020). *Premarket Notification 510(k) 03/13*. https://www.fda.gov/medical-devices

FDA. (2021). *Catalog of Regulatory Science Tools to Help Assess New Medical Devices. 03/23*. https://www.fda.gov/medical-devices

Federal Ministry of Justice and Consumer Protection. (2019, November 20). *Federal Data Protection Act of 30 June 2017 (Federal Law Gazette I p. 2097), as last amended by Article 12 of the Act of 20 November 2019 (Federal Law Gazette I, p. 1626)*. https://www.gesetze-im-internet.de/englisch_bdsg/

Federal Trade Commission. (2020). *Privacy & security update for 2019*. https://www.ftc.gov/reports/privacy-data-security-update-2019

Feiner, L., & Graham, M. (2020, June 16). *Pelosi says advertisers should use their "tremendous leverage" to force social media companies to stop spreading false and dangerous information*. CNBC. https://www.cnbc.com/2020/06/16/pelosi-says-advertisers-should-push-platforms-to-combat-disinformation.html

Feldman, J. (2020, October 18). *Flickplay's 3d social media platform presents as an industry first*. Influencive. https://www.influencive.com/flickplays-3d-social-media-platform-presents-as-an-industry-first/

Fernández-Macías, E., Hurley, J., Peruffo, E., Storrie, D., Poel, M., & Packalén, E. (2018). *Game-changing technologies: Exploring the impact of production processes and work*. Eurofound. https://www.eurofound.europa.eu/publications/report/2018/game-changing-technologies-in-european-manufacturing

Fieser, J. (2021). David Hume (1711—1776). In *The internet encyclopedia of philosophy*. https://iep.utm.edu/hume/

Fink, C., & Kenny, C. J. (2003). W (h) ither the digital divide? *Info, 5*(6), 15-24.

Floridi, L. (2018, February 17). *Soft ethics and the governance of the Digital.* Philosophy & Technology. https://link.springer.com/article/10.1007/s13347-018-0303-9

Floridi, L. (2021). The European legislation on AI: A brief analysis of its philosophical approach. *Philosophy & Technology*, *34*(2), 215–222. doi:10.100713347-021-00460-9 PMID:34104628

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions - Royal Society. Mathematical, Physical, and Engineering Sciences*, *374*(2083), 20160360. Advance online publication. doi:10.1098/rsta.2016.0360 PMID:28336805

Foer, F. (2017, September 19). Facebook's war on free will. *The Guardian*. https://www.theguardian.com/technology/2017/sep/19/facebooks-war-on-free-will

Foley, D. (2016, March 9). *Here's the original 'Two minutes hate'*. Intellectual Takeout. https://www.intellectualtakeout.org/Blog/Heres-Original-Two-Minutes-Hate/

Food and Drug Administration & Department of Health and Human Services. (1997). Title 21 Code of Federal Regulations Part 11: Electronic Records; Electronic Signatures. *Federal Register*, *62*, 13464.

Food and Drug Administration. (2017, November 13). *FDA approves pill with sensor that digitally tracks if patients have ingested their medication.* FDA News Release. https://www.fda.gov/news-events/press-announcements/fda-approves-pill-sensor-digitally-tracks-if-patients-have-ingested-their-medication

Forbes Technology Council. (2021, January 6). Tech experts predict 13 areas AI and VR are set to revolutionize. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2021/01/06/tech-experts-predict-13-areas-ai-and-vr-are-set-to-revolutionize/?sh=3cd7283b25b0

Ford, M. (2018). *Architects of intelligence*. Packt Publishing.

Foth, M., Anastasiu, I., Mann, M., & Mitchell, P. (2021). From Automation to Autonomy: Technological Sovereignty for Better Data Care in Smart Cities. In B. T. Wang & C. M. Wang (Eds.), *Automating Cities. Advances in 21st Century Human Settlements*. Springer.

Freeman, M. (2020). *A Magna Carta for Children? Rethinking Children's Rights*. Cambridge University Press.

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. MIT Press. doi:10.7551/mitpress/7585.001.0001

Furey, H., & Martin, F. (2019). AI Education Matters: A Modular Approach to AI Ethics Education. *AI Matters: A quarterly newsletter of the ACM Special Interest Group in Artificial Intelligence*.

Gal, U., Jensen, T. B., & Stein, M.-K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, *30*(2), 100301. Advance online publication. doi:10.1016/j.infoandorg.2020.100301

Garrahan, M. (2009, December 4). The rise and fall of Myspace. *Financial Times*. https://www.ft.com/content/fd9ffd9c-dee5-11de-adff-00144feab49a

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2018). *Datasheets for datasets*. https://arxiv.org/abs/1803.09010v7

General Data Protection Regulations. (2018). https://gdpr-info.eu

Gewirtz, D. (2018, March 18). *Volume, velocity, and variety: Understanding the three V's of big data*. ZDNet. https://www.zdnet.com/article/volume-velocity-and-variety-understanding-the-three-vs-of-big-data

Gilshan, D. (2021). *The ethics of diversity*. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2021/02/03/the-ethics-of-diversity/

Global social media stats. (2021). *Datareportal*. https://datareportal.com/social-media-users

Gomes, R. (2020). *Pope: Church's social teaching can help AI serve the common good*. Vatican News. https://www.vaticannews.va/en/pope/news/2020-02/pope-francis-artificial-intelligence-algor-ethics.html

Gonçalves, P. (2019, January 15). *HSBC banks on Blockchain tech to process $250bn worth of FX transactions*. International Investment. https://www.internationalinvestment.net/news/4000471/hsbc-banks-blockchain-tech-process-usd250bn-worth-transactions

Goodell, G., & Aste, T. (2019). Can cryptocurrencies preserve privacy and comply with regulations? *Frontiers in Blockchain*, *2*, 4. Advance online publication. doi:10.3389/fbloc.2019.00004

Goodson, S. (2012, March 5). If you're not paying for it, you become the product. *Forbes*. https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/?sh=3c309c6e5d6e

Gosselin, A. (2020). At Home in a Psychiatric Hospital. *Social Philosophy Today*, *36*(July), 71–87. doi:10.5840ocphiltoday202012971

Gottlieb, S. (2018). *Digital Health Innovation Action Plan*. https://www.fda.gov/media/106331/

Gottlieb, S. (2018). *Transforming FDA's Approach to Digital Health*. https://www.fda.gov/news-events/speeches-fda-officials/transforming-fdas-approach-digital-health-04262018

Goyal, M., Ospel, J. M., Ganesh, A., Marko, M., & Fisher, M. (2021). Rethinking consent for stroke trials in time-sensitive situations: Insights from the COVID-19 pandemic. *Stroke*, *52*(4), 1527–1531. doi:10.1161/STROKEAHA.120.031976 PMID:33588599

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, *62*, 729–754. doi:10.1613/jair.1.11222

Green, B., & Kak, A. (2021, June 15). The false comfort of human oversight as an antidote to A.I. harm. *Slate*. https://slate.com/technology/2021/06/human-oversight-artificial-intelligence-laws.html?via=rss_socialflow_twitter

Gupta, K. (2019). *Medical Entanglements: Rethinking Feminist Debates about Healthcare*. Rutgers University Press.

Hadden, K. B., Prince, L. Y., Moore, T. D., James, L. P., Holland, J. R., & Trudeau, C. R. (2017). Improving readability of informed consents for research at an academic medical institution. *Journal of Clinical and Translational Science*, *1*(6), 361–365. doi:10.1017/cts.2017.312 PMID:29707258

Hallinan, D., & Friedewald, M. (2015). Open consent, biobanking and data protection law: Can open consent be "informed" under the forthcoming data protection regulation? *Life Sciences, Society and Policy*, *11*(1), 1. Advance online publication. doi:10.118640504-014-0020-9 PMID:26085311

Hamington, M. (2019). Integrating Care Ethics and Design Thinking. *Journal of Business Ethics*, *155*(1), 91–103. doi:10.100710551-017-3522-6

Haraway, D. J. (2013). *When Species Meet*. University of Minnesota Press.

Hari, J. (2018). *Lost Connections: Uncovering the Real Causes of Depression and the Unexpected Solutions*. Bloomsbury Publishing USA.

Harmon, A. (1996). Daily Life's Digital Divide. *Los Angeles Times, 3*, A1. Retrieved May 10, 2021, from http://articles.latimes.com/1996-07-03/news/mn-20785_1_digital-technology

Harris, B. (2017). Uber, Lyft, and regulating the sharing economy. *Seattle University Law Review*, *41*(1), 269–285. https://digitalcommons.law.seattleu.edu/sulr/vol41/iss1/8/

Hawkins, B. L., & Oblinge, D. G. (2006). The Myth about the Digital Divide. *EDUCAUSE Review*, *41*(4), 12–13.

Hazel, J. W., & Clayton, E. W. (2021). *Law Enforcement and Genetic Data*. Bioethics Briefings. The Hasting Center. https://www.thehastingscenter.org/briefingbook

Heilweil, R. (2020, February 28). *The Pope's plan to fight back against evil AI*. Vox. https://www.vox.com/recode/2020/2/28/21157760/pope-vatican-artificial-intelligence

Hendricks, S. (2018, June 14). *Dark forest theory: A terrifying explanation of why we haven't heard from aliens yet*. Big Think. https://bigthink.com/scotty-hendricks/the-dark-forest-theory-a-terrifying-explanation-of-why-we-havent-heard-from-aliens-yet

Hess, J. L., & Fore, G. (2018). A systematic literature review of US engineering ethics interventions. *Science and Engineering Ethics*, *24*(2), 551–583. doi:10.100711948-017-9910-6 PMID:28401510

High-Level Expert Group on AI. (2019). *Ethics guidelines for trustworthy AI*. Available in https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Hilbert, M. (2010, August). *The manifold definitions of the digital divide and their diverse implications for policy responsibility*. TPRC.

Hilbert, M. (2011, December). Digital gender divide or technologically empowered women in developing countries? A typical case of lies, damned lies, and statistics. *Women's Studies International Forum*, *34*(6), 479–489. doi:10.1016/j.wsif.2011.07.001

Hill, K. (2020, March 18). The secretive cimpany that might end privacy as we know it. *New York Times*. https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html

Hilty, D. M., Ferrer, D. C., Parish, M. B., Johnston, B., Callahan, E. J., & Yellowlees, P. M. (2013). The Effectiveness of Telemental Health: A 2013 Review. *Telemedicine Journal and E-Health, 19*(6), 444–454.

Hochheiser, H., & Valdez, R. S. (2020). *Human-Computer interaction, ethics, and Biomedical Informatics*. Yearbook of medical informatics. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7442500/#:~:text=Ethics%20in%20Human%2DComputer%20Interaction,and%20accountability%2C%20among%20others%201%20

Holm, S. (2019). Bioethics and mental health—An uneasy relationship. *Ethics*. *Medicine and Public Health*, *10*, 1–7.

Hou, J. U., Kim, D., Ahn, W. H., & Lee, H. K. (2018). Copyright Protections of Digital Content in the Age of 3D Printer: Emerging Issues and Survey. *IEEE Access: Practical Innovations, Open Solutions*, *6*, 44082–44093. doi:10.1109/ACCESS.2018.2864331

Hsieh, Y.-Y., Vergne, J.-P., Anderson, P., Lakhani, K., & Reitzig, M. (2018). Bitcoin and the rise of decentralized autonomous organizations. *Journal of Organization Design*, *7*(14), 14. Advance online publication. doi:10.118641469-018-0038-1

Hu, P., Dhelim, S., Ning, H., & Qiu, T. (2017). Survey on fog computing: architecture, key technologies, applications and open issues. Journal of Network and Computer Applications. doi:10.1016/j.jnca.2017.09.002

Huang, Q., Ma, Z., Fu, J., Niu, X., & Yang, Y. (2013). *Attribute Based DRM Scheme with Efficient Revocation in Cloud Computing* (Vol. 8). doi:10.4304/jcp.8.11.2776-2781

Huang, J., Lu, P., Juang, W., Fan, C., Lin, Z., & Lin, C. (2014). Secure and efficient digital rights management mechanisms with privacy protection. *Journal of Shanghai Jiaotong University (Science)*, *19*(4), 443–447. doi:10.100712204-014-1523-5

Huang, Q., Fu, J., Ma, Z., Yang, Y., & Niu, X. (2014). Encrypted data sharing with multi-owner based on digital rights management in online social networks. *Journal of China Universities of Posts and Telecommunications*, *21*(1), 86–93. doi:10.1016/S1005-8885(14)60273-9

Hume, D. (1985). *A treatise of human nature*. Penguin Classics. (Original work published 1740)

Hussain, S. A., Fatima, M., Saeed, A., Raza, I., & Shahzad, R. K. (2017). Multilevel classification of security concerns in cloud computing. *Applied Computing and Informatics, 13*(1), 57–65.

Hyrynsalmi, S., Hyrynsalmi, S. M., & Kimppa, K. K. (2020). Blockchain ethics: A systematic literature review of Blockchain research. In M. Cacace, R. Halonen, H. Li, T. P. Orrensalo, C. Li, G. Widén, & R. Suomi (Eds.), *Well-Being in the information society. Fruits of respect* (Vol. 1270, pp. 145–155). Springer. doi:10.1007/978-3-030-57847-3_10

Iftikhar, S., Kamran, M., Munir, E. U., & Khan, S. U. (2017). A Reversible Watermarking Technique for Social Network Data Sets for Enabling Data Trust in Cyber, Physical, and Social Computing. *IEEE Systems Journal*, *11*(1), 197–206. doi:10.1109/JSYST.2015.2416131

Illing, S. (2018, April 4). Cambridge Analytica, the shady data firm that might be a key Trump-Russia link, explained. *Vox*. https://www.vox.com/policy-and-politics/2017/10/16/15657512/cambridge-analytica-facebook-alexander-nix-christopher-wylie

Institute of Electrical and Electronics Engineers. (2021). *The IEEE global initiative on ethics of extended reality*. IEEE Standards Association. https://standards.ieee.org/industry-connections/ethics-extended-reality.html

International Swaps and Derivatives Association. (2017). *Smart contracts and distributed ledger - A legal perspective* [White paper]. https://www.isda.org/a/6EKDE/smart-contracts-and-distributed-ledger-a-legal-perspective.pdf

Internet Live Stats. (2021). Retrieved May 10, 2021, from https://www.internetlivestats.com/internet-users/

Internet World Stats. (2021). *Usage and Population Statistics*. Retrieved May 10, 2021, from https://www.internetworldstats.com/stats.htm

Iserson, K.V., & Chiasson, P.M. (2002). The ethics of applying new medical technologies. *Semin Laparosc Surg., 9*(4), 222-9. doi:10.1053/slas.2002.36465

Islam, G., & Greenwood, M. (2021). Reconnecting to the social in business ethics. *Journal of Business Ethics*. https://link.springer.com/article/10.1007/s10551-021-04775-7

Jaeger, J. (2021, March 19). Popular Clubhouse app being probed for GDPR violations. *Compliance Week*. https://www.complianceweek.com/gdpr/popular-clubhouse-app-being-probed-for-gdpr-violations/30181.article

Jago, A. S. (2019). Algorithms and authenticity. *Academy of Management Discoveries*, *5*(1), 38–56. doi:10.5465/amd.2017.0002

Jain, R. (2020b, Dec 21). *Payments strategies and business model innovation (1.0)* [Online course]. Centre for Financial Technology and Entrepreneurship. https://courses.cfte.education/payment-strategies-business-model-innovations-course

Jain, R., Mcloughlin, C., Bar, F., & Ford, K. (2020, Dec 21). *Payments in digital finance 1.0* [Online course]. Centre for Financial Technology and Entrepreneurship. https://courses.cfte.education/payments-in-digital-finance-specialisation

Jain. (2020a, December 15). *Fintech Future.* SPD Salford Professional Development - Gov Tech UK.

Jarrahi, M. H., & Sutherland, W. (2019). Algorithmic management and algorithmic competencies: Understanding and appropriating algorithms in Gig Work. In N. Taylor, C. Christian-Lamb, M. Martin, & B. Nardi (Eds.), *Information in contemporary society* (pp. 578–589). Springer. doi:10.1007/978-3-030-15742-5_55

Jawad, A. J. (2021). Bioethics of medical devices based on brain computer interfaces (BCI). *Journal of Clinical Research & Bioethics*. https://www.academia.edu/45585122/Bioethics_of_Medical_Devices_Based_on_Brain_Computer_Interfaces_BCI_

Jayasinghe, N., Moallem, B. I., Kakoullis, M., Ojie, M.-J., Sar-Graycar, L., Wyka, K., Reid, M. C., & Leonard, J. P. (2018). Establishing the feasibility of a tablet-based consent process with older adults: A mixed-methods study. *The Gerontologist*, *59*(1), 124–134. doi:10.1093/geront/gny045 PMID:29757375

Jenkins, H. (2006). *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. MacArthur Foundation.

Jeong, S. (2017, April 4). Mastodon is like Twitter without nazis, so why are we not using it? *Vice*. https://www.vice.com/en/article/783akg/mastodon-is-like-twitter-without-nazis-so-why-are-we-not-using-it

Jermy, S., & Peng, L. (2018). *Assurance over machine learning and algorithms: Trust in the age of algorithms, robots, and cognitive technologies*. Deloitte. https://www2.deloitte.com/au/en/pages/audit/articles/assurance-over-machine-learning-algorithms.html

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. doi:10.103842256-019-0088-2

John, M. (2003). *Children's Rights and Power: Charging Up for a New Century*. Jessica Kingsley Publishers.

Johns Hopkins University. (2021). *COVID-19 dashboard*. Retrieved August 3, 2021. https://coronavirus.jhu.edu/map.html

Joshi, N., & Petrlic, R. (2013). Towards practical privacy-preserving digital rights management for cloud computing. *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC),* 265–270. 10.1109/CCNC.2013.6488456

Joshua. (2020, January 11). *Comprehensive list of banks using Blockchain technology in 2020*. Hackernoon. https://hackernoon.com/comprehensive-list-of-banks-using-blockchain-technology-in-2020-revised-and-updated-uq493yrb

Jowett, G., & O'Donnell, V. (2006). *Propaganda and persuasion* (4th ed.). Sage.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454. doi:10.1016/0010-0285(72)90016-3

Kak, S. (2019, October 16). *Why a computer will never be truly conscious*. The Conversation. https://theconversation.com/why-a-computer-will-never-be-truly-conscious-120644

Kakarlapudi, P. V., & Mahmoud, Q. H. (2021). A systematic review of blockchain for consent management. *Healthcare (Basel, Switzerland)*, *9*(2), 137. Advance online publication. doi:10.3390/healthcare9020137 PMID:33535465

Kalifa, R. (2021). *Kalifa review of UK Fintech*. UK Treasury. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/978396/KalifaReviewofUKFintech01.pdf

Kant, I. (1785). Grundlegung zur Metaphysik der Sitten, 1785. In *Immanuel Kant, Gesamtausgabe in zehn Bänden, Leipzig 1838, Modes und Baumann* (p. 23). Academic Press.

Kant, I. (1929). *Critique of pure reason* (7th ed.).

Kaplan, B. (2021). *Regulation of Software as a Medical Device: Opportunity for Bioethics*. Hastings Bioethics Forum. https://www.thehastingscenter.org/

Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2014). Dynamic consent: A patient interface for twenty-first century research networks. *European Journal of Human Genetics*, *23*(2), 141–146. doi:10.1038/ejhg.2014.71 PMID:24801761

Kearns, M., & Roth, A. (2020). *Ethical algorithm design should guide technology regulation*. Brookings Institution's Artificial Intelligence and Emerging Technology Initiative. https://www.brookings.edu/research/ethical-algorithm-design-should-guide-technology-regulation/

Keirns, G. (2017, March 30). *Major banks, start-ups advance syndicated loan pilots*. Coindesk. https://www.coindesk.com/banks-startups-Blockchain-syndicated-loans

Kemp, S. (2020, January 30). *Digital 2020: 3.8 billion people use social media*. We Are Social. https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media#

Khanzode, K. C. A., & Sarode, R. D. (2020). Advantages and disadvantages of artificial intelligence and machine learning: A literature review. *International Journal of Library and Information Science*, *9*(1), 30–36.

Khozin, S. (2017). Information Exchange and Data Transformation (INFORMED) An integrated approach to big data analytics. *Nat Rev Drug Discov*. https://www.ehidc.org/

Kieckens, E. (2020, November 6). *Pope embraces AI and robotics that serve the common good*. Innovation Origins. https://innovationorigins.com/pope-embraces-ai-and-robotics-that-serve-the-common-good/

Kimbrell, G. (2018). *Is Regulation Killing Innovation in Health Care?* Forbes Technology Council. https://www.forbes.com/sites/forbestechcouncil

Kinni, T. (2016). People management by algorithm: What's happening this week at the intersection of management and technology. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/tech-savvy-people-management-by-algorithm/

Kishigami, J., Fujimura, S., Watanabe, H., Nakadaira, A., & Akutsu, A. (2015). The Blockchain-based Digital Content Distribution System. *Proceedings 2015 Ieee Fifth International Conference on Big Data and Cloud Computing Bdcloud 2015*, 187–190. 10.1109/BDCloud.2015.60

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *The Journal of Legal Analysis*, *10*, 113–174. doi:10.1093/jla/laz001

Kluger, J. (2020, August 27). Online Therapy, Booming During the Coronavirus Pandemic, May Be Here to Stay. *Time*. https://time.com/5883704/teletherapy-coronavirus/

Koch, C. (2018). What is consciousness? *Scientific American*, *318*(6), 60–64. doi:10.1038cientificamerican0618-60 PMID:29949559

Koch, C. (2019). Proust among the machines. *Scientific American*, *321*(6), 46–49. doi:10.1038/10.1038cientificamerican1219-46

Koulouzis, S., Mousa, R., Karakannas, A., de Laat, C., & Zhao, Z. (2018). Information Centric Networking for Sharing and Accessing Digital Objects with Persistent Identifiers on Data Infrastructures. *Proceedings of the 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 661–668. 10.1109/CCGRID.2018.00098

Kount. (2019, June 26). *Kount launches next-generation AI, changing how payments fraud prevention is delivered*. https://kount.com/announcements/kount-launches-next-generation-ai/

Kraft, S. A., Garrison, N. A., & Wilfond, B. S. (2019). Understanding as an ethical aspiration in an era of digital technology-based communication: An analysis of informed consent functions. *The American Journal of Bioethics*, *19*(5), 34–36. doi:10.1080/15265161.2019.1587035 PMID:31090520

Krishna, D., Albinson, N., & Chu, Y. (2017). *Managing algorithmic risks: Safeguarding the use of complex algorithms and machine learning*. Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/risk/us-risk-algorithmic-machine-learning-risk-management.pdf

Kritikos, M. (2018). *What if algorithms could abide by ethical principles?* European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/ATAG/2018/624267/EPRS_ATA(2018)624267_EN.pdf

Kularski, C., & Moller, S. (2012). The digital divide as a continuation of traditional systems of inequality. *Sociology*, *5151*, 1–23.

Kumar, S. (2019, November 25). *Advantages and disadvantages of artificial intelligence.* Towards Data Science. https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-182a5ef6588c

Kwon, G. R., Lama, R. K., Pyun, J. Y., & Park, C. S. (2016). Multimedia digital rights management based on selective encryption for flexible business model. *Multimedia Tools and Applications*, *75*(12), 6697–6715. doi:10.100711042-015-2563-z

Lai, J., & Widmar, N. O. (2021). Revisiting the Digital Divide in the COVID-19 Era. *Applied Economic Perspectives and Policy*, *43*(1), 458–464. doi:10.1002/aepp.13104 PMID:33230409

Langarizadeh, M., Tabatabaei, M. S., Tavakol, K., Naghipour, M., Rostami, A., & Moghbeli, F. (2017). Telemental Health Care, an Effective Alternative to Conventional Mental Care: A Systematic Review. *Acta Informatica Medica*, *25*(4), 240–246. doi:10.5455/aim.2017.25.240-246 PMID:29284913

Lankton, N. K., McKnight, D. H., & Tripp, J. F. (2017). Facebook privacy management strategies: A cluster analysis of user privacy behaviors. *Computers in Human Behavior*, *76*, 149–163. doi:10.1016/j.chb.2017.07.015

Lapointe, C., & Fishbane, L. (2019). The Blockchain ethical design framework. *Innovations: Technology, Governance, Globalization*, *12*(3–4), 50–71. doi:10.1162/inov_a_00275

Lauritsen, K. J., & Nguyen, T. (2009). Combination Products Regulation at the FDA. *Clinical Pharmacology and Therapeutics*, *85*(5), 468–470. doi:10.1038/clpt.2009.28 PMID:19381151

Le Breton, S., Lamberti, M. J., Dion, A., & Getz, K. A. (2020, October 22). *COVID-19 and Its impact on the future of clinical trial execution*. Applied Clinical Trials. https://www.appliedclinicaltrialsonline.com/view/covid-19-and-its-impact-on-the-future-of-clinical-trial-execution

Lee, J. (2015, July 21). *Retailer forever 21 deploys M2SYS technology's biometric time clock*. Biometric Update. https://www.biometricupdate.com/201507/retailer-forever-21-deploys-m2sys-technologys-biometric-time-clock

Lee, Y. N. (2017, December 18). *Robots 'are here to give us a promotion,' not take away jobs, Gartner says*. CNBC. https://www.cnbc.com/2017/12/18/artificial-intelligence-will-create-more-jobs-than-it-ends-gartner.html

Lee, C. C., Li, C. T., Chen, Z. W., Lai, Y. M., & Shieh, J. C. (2018). An improved E-DRM scheme for mobile environments. *Journal of Information Security and Applications*, *39*, 19–30. doi:10.1016/j.jisa.2018.02.001

Lee, H., Park, S., Seo, C., & Shin, S. U. (2016a). DRM cloud framework to support heterogeneous digital rights management systems. *Multimedia Tools and Applications*, *75*(22), 14089–14109. doi:10.100711042-015-2662-x

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1). Advance online publication. doi:10.1177/2053951718756684

Legault, G. A., Béland, J. P., Parent, M., Bédard, S. K., Bellemare, C. A., Bernier, L., Dagenais, P., Daniel, C. E., Gagnon, H., & Patenaude, J. (2019). Ethical Evaluation in Health Technology Assessment: A Challenge for Applied Philosophy. *Open Journal of Philosophy*, *9*, 331–351. doi:10.4236/ojpp.2019.93022

Lepri, B., Staiano, J., Sangokoya, D., Letouzé, E., & Oliver, N. (2017). The tyranny of data? The bright and dark sides of data-driven decision-making for social good. In T. Cerquitelli, D. Quercia, & F. Pasquale (Eds.), *Transparent data mining for big and small data* (Vol. 32, pp. 3–24). Springer. doi:10.1007/978-3-319-54024-5_1

Lernende Systeme. (2019). *Work, training and human-machine interaction* [White paper]. Future of Work and Human-Machine Interaction Working Group. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen_EN/AG2_Whitepaper_Executive_Summary_final_200204.pdf

Lidz, C. W., Albert, K., Appelbaum, P., Dunn, L. B., Overton, E., & Pivovarova, E. (2015). Why is therapeutic misconception so prevalent? *Cambridge Quarterly of Healthcare Ethics*, *24*(2), 231–241. doi:10.1017/S096318011400053X PMID:25719358

Li, G., Yin, C., Zhou, Y., Wang, T., Chen, J., Liu, Y., Chen, T., Wang, H., Zhang, L., & Chen, X. (2020). Digitalized adaptation of oncology trials during and after COVID-19. *Cancer Cell*, *38*(2), 148–149. doi:10.1016/j.ccell.2020.06.018 PMID:32634378

Lindgren, S. (2017). *Digital Media and Society*. Sage.

Linklaters. (2019). *Leadership, governance, systems and controls*. https://www.linklaters.com/en/insights/publications/2019/february/ethics-in-banking-and-finance/leadership-governance-systems-and-controls

Lin, P. (2016). Why ethics matters for autonomous cars. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous driving: Technical, legal and social aspects* (pp. 69–85). Springer. doi:10.1007/978-3-662-48847-8_4

Lin, X.-J., Sun, L., Qu, H., & Zhang, X. (2021). Public key encryption supporting equality test and flexible authorization without bilinear pairings. *Computer Communications*, *170*, 190–199. doi:10.1016/j.comcom.2021.02.006

Lin, X.-J., Wang, Q., Sun, L., & Qu, H. (2021). Identity-based encryption with equality test and datestamp-based authorization mechanism. *Theoretical Computer Science*, *861*, 117–132. doi:10.1016/j.tcs.2021.02.015

Litwin, J. (2016). Engagement shift: Informed consent in the digital era: Why electronic informed consent is key to supporting today's patient-centric mantra in clinical trials. *Applied Clinical Trials*, *25*(6-7). https://www.appliedclinicaltrialsonline.com/view/engagement-shift-informed-consent-digital-era-0

Livingstone, S., Lansdown, G., & Third, A. (2017). *The Case for a UNCRC General Comment on Children's Rights and Digital Media.* LSE Consulting. https://www.childrenscommissioner.gov.uk/report/the-case-for-a-uncrc-general-comment-on-childrens-rights-and-digital-media/

Livingstone, S. (2014). Children's Digital Rights: A Priority. *LSE Research Online.*, *42*(4/5), 20–24.

Livingstone, S. (2018). Children: A Special Case for Privacy? *Intermedia.*, *46*(2), 18–23.

Loggins, K. (2020, February 24). *Here's what happens when an algorithm determines your work schedule*. Vice. https://www.vice.com/en_us/article/g5xwby/heres-what-happens-when-an-algorithm-determines-your-work-schedule

London, A. J., & Danks, D. (2018). Regulating autonomous vehicles: A policy proposal. In *Proceedings of the 2018 AAAI/ACM Conference on artificial intelligence, ethics, and society* (pp. 216-221). Association for Computing Machinery. 10.1145/3278721.3278763

Lopez, P. (2019, July 23). *Five principles of ethics to help drive Fintech platform risk management*. Linkedin. https://www.linkedin.com/pulse/five-principles-ethics-help-drive-fintech-platform-risk-lopez/

Lundy, L., Kilkelly, U., & Byrne, B. (2013). Incorporation of the United Nations Convention on the Rights of the Child in Law: A Comparative Review. *International Journal of Children's Rights*, *21*(3), 442–463. doi:10.1163/15718182-55680028

Lurie, J., & Tjelflaat, T. (2012). Children's Rights and the UN Convention on the Rights of the Child: Monitoring in Norway. *Dialogue in Praxis, 1*(14), 41-56.

Lütge, C., & Uhl, M. (2021). *Business ethics: Interdisciplinary perspectives and future challenges*. Oxford Scholarship Online. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780198864776.001.0001/oso-9780198864776-chapter-8

Lu, X., Pan, Z., & Xian, H. (2020). An integrity verification scheme of cloud storage for internet-of-things mobile terminal devices. *Computers & Security*, *92*, 101686. doi:10.1016/j.cose.2019.101686

MacKay, D., & Danis, M. (2016). Federalism and Responsibility for Health Care. *Public Affairs Quarterly*, *30*(1), 1–29.

Mackhight, J. (2019, April 1). *An ethical framework for the AI age*. The Banker. https://www.thebanker.com/Transactions-Technology/An-ethical-framework-for-the-AI-age

Maguire, A. (2020). *The difference between the terms equality, equity, and liberation, illustrated*. Interaction Institute for Social Change.

Mahler, M., Auza, C., Albesa, R., Melus, C., & Wu, J. A. (2021). Regulatory aspects of artificial intelligence and machine learning-enabled software as medical devices (SaMD). doi:10.1016/B978-0-12-820239-5.00010-

Makepeace, N. (n.d.). Data ethics starts at the top. *GWI*. https://gwi.com.au/blog/data-ethics-starts-at-the-top/

Marckmann, G., & Goodman, K. (2006). Introduction: Ethics of Information Technology in Health Care. *International Journal of Information Ethics*, *5*, 2–5. doi:10.29173/irie188

Marr, B. (2020). *What are the negative impacts of artificial intelligence (AI)?* https://bernardmarr.com/default.asp?contentID=1827

Martinho, A., Kroesen, M., & Chorus, C. (2021). Computer says I don't know: An empirical approach to capture moral uncertainty in artificial intelligence. *Minds and Machines*, *31*(2), 215–237. doi:10.100711023-021-09556-9

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. doi:10.100710551-018-3921-3

Mateescu, A., & Nguyen, A. (2019, February 6). Algorithmic management in the workplace. *Data & Society*. https://datasociety.net/library/explainer-algorithmic-management-in-the-workplace/

Mathiesen, K. (2012) *Human rights for the Digital age*. Taylor & Francis. https://www.tandfonline.com/doi/abs/10.1080/08900523.2014.863124

Mayor, A. (2018). Gods and Robots. In Myths, Machines and Ancient Dreams of Technology. Princeton University Press. doi:10.2307/j.ctvc779xn

Ma, Z. F., Huang, J. Q., Jiang, M., & Niu, X. X. (2016). A Novel Image Digital Rights Management Scheme with High-Level Security, Usage Control and Traceability. *Chinese Journal of Electronics*, *25*(3), 481–494. doi:10.1049/cje.2016.05.014

Ma, Z. F., Huang, W. H., Bi, W., Gao, H. M., & Wang, Z. (2018). A Master-Slave Blockchain Paradigm and Application in Digital Rights Management. *China Communications*, *15*(8), 174–188. doi:10.1109/CC.2018.8438282

Ma, Z. F., Jiang, M., Gao, H. M., & Wang, Z. (2018). Blockchain for digital rights management. *Future Generation Computer Systems-the International Journal of Escience*, *89*, 746–764. doi:10.1016/j.future.2018.07.029

McGraw, D., Greene, S. M., Miner, C. S., Staman, K. L., Welch, M. J., & Rubel, A. (2015). Privacy and confidentiality in pragmatic clinical trials. *Clinical Trials*, *12*(5), 520–529. doi:10.1177/1740774515597677 PMID:26374682

McQuillan, D. (2018). People's councils for ethical machine learning. *Social Media + Society*, *4*(2). Advance online publication. doi:10.1177/2056305118768303

Menghwar, P. S., & Daood, A. (2021). Creating shared value: A systematic review, synthesis and integrative perspective. *International Journal of Management Reviews*, ijmr.12252. Advance online publication. doi:10.1111/ijmr.12252

Microsoft. (2020, March 3). *Digital campaign launches to give young people the algorithm literacy they deserve*. https://news.microsoft.com/en-ca/2020/03/03/digital-campaign-launches-to-give-young-people-the-algorithm-literacy-they-deserve/

Millar, A. P., Adeyemi, O., Behrmann, G., Fuhrmann, P., Garonne, V., Litvinsev, D., Mkrtchyan, T., Rossi, A., Sahakyan, M., & Starek, J. (2018). Storage for Advanced Scientific Use-Cases and Beyond. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP),* 651–657. 10.1109/PDP2018.2018.00109

Millar, J., Lin, P., Abney, K., & Bekey, G. (2017). Ethics settings for autonomous vehicles. In L. Patrick, R. Jenkins, & K. Abney (Eds.), *Robot ethics 2.0: From autonomous cars to artificial intelligence* (pp. 20–34). Oxford University Press.

Mills, K. A., Stornaiuolo, A., Smith, A., & Jessica Zacher, P. (Eds.). (2017). *Handbook of Writing, Literacies, and Education in Digital Cultures* (1st ed.). Routledge. doi:10.4324/9781315465258

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the 2019 Conference on fairness, accountability, and transparency* (pp. 220-229). Association for Computing Machinery. 10.1145/3287560.3287596

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). Advance online publication. doi:10.1177/2053951716679679

Möhlmann, M., & Henfriendon, O. (2019, August 30). What people hate about being managed by algorithms, according to a study of Uber drivers. *Harvard Business Review*. https://hbr.org/2019/08/what-people-hate-about-being-managed-by-algorithms-according-to-a-study-of-uber-drivers

Mokyr, J., Vickers, C., & Ziebarth, N. L. (2015). The history of technological anxiety and the future of economic growth: Is this time different? *The Journal of Economic Perspectives*, *29*(3), 31–50. doi:10.1257/jep.29.3.31

Mol, A. (2008). *The Logic of Care: Health and the Problem of Patient Choice*. Routledge. doi:10.4324/9780203927076

Montague, E., Day, T. E., Barry, D., Brumm, M., McAdie, A., Cooper, A. B., Wignall, J., Erdman, S., Núñez, D., Diekema, D., & Danks, D. (2021). The case for information fiduciaries: The implementation of a data ethics checklist at Seattle Children's Hospital. *Journal of the American Medical Informatics Association: JAMIA*, *28*(3), 650–652. doi:10.1093/jamia/ocaa307 PMID:33404593

Moor, J. H. (1985). What is computer ethics? *Metaphilosophy*, *16*(4), 266–275. doi:10.1111/j.1467-9973.1985.tb00173.x

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, *21*(4), 18–21. doi:10.1109/MIS.2006.80

Moran, W. (2021) *Technical career Institutes, Inc., (TCI College) strategic analysis*. CUNY Academic Works. https://academicworks.cuny.edu/qb_pubs/72/

Morley, J., Widdicks, K., & Hazas, M. (2018). Digitalisation, energy, and data demand: The impact of Internet traffic on overall and peak electricity consumption. *Energy Research & Social Science*, *38*, 128–137. doi:10.1016/j.erss.2018.01.018

Morrison, K. R., & Matthes, J. (2011). Socially motivated projection: Need to belong increases perceived opinion consensus on important issues. *European Journal of Social Psychology*, *41*(6), 707–719. doi:10.1002/ejsp.797

Morrison, R., Mazey, N. C., & Wingreen, S. C. (2020). The DAO controversy: The case for a new species of corporate governance? *Frontiers in Blockchain*, *3*, 25. Advance online publication. doi:10.3389/fbloc.2020.00025

Mtech, R. K. (2015). *The Non-Tangible Masking of Confidential Information using Video Steganography*. Academic Press.

Mulder, T., & Tudorica, M. (2019). Privacy policies, cross-border health data and the GDPR. *Information & Communications Technology Law*, *28*(3), 261–274. doi:10.1080/13600834.2019.1644068

Müller, V. C. (2020). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020 ed.). Stanford University. https://plato.stanford.edu/archives/win2020/entries/ethics-ai/

Munier, M., Lalanne, V., & Ricarde, M. (2012). Self-Protecting Documents for Cloud Storage Security. *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, 1231–1238. 10.1109/TrustCom.2012.261

Nalini, B. (2019, May 1). *The hitchhiker's guide to AI ethics.* Towards Data Science. https://towardsdatascience.com/ethics-of-ai-a-comprehensive-primer-1bfd039124b0

Narayan, A., & Felten, E. W. (2014). *No silver bullet: De-identification still doesn't work*. Princeton University. https://www.cs.princeton.edu/~arvindn/publications/no-silver-bullet-de-identification.pdf

NASA authorization 1965. hearings before the Committee on Science and Astronautics, U.S. House of Representatives, Eighty-eighth Congress, second session, on H. R. 9641, superseded by H. R. 10456. Washington: GovtU. S. https://catalog.hathitrust.org/Record/100718937/Cite

Nash, D. (2021). *Healthcare Technology's Digital Dilemma. Should digital connectivity be considered a "vital sign"?* https://www.medpagetoday.com

National Information Center on Health Services Research & Health Care Technology (NICHSR). (2021). *National Library of Medicine*. https://www.nlm.nih.gov/

National Public Radio. (2011, February 24). *The Supreme court's failure to protect Blacks' rights.* https://www.npr.org/2011/02/24/133960082/the-supreme-courts-failure-to-protect-civil-rights

Nersessian, D. (2018). *The law and Ethics of Big Data Analytics: A new role for international human rights in the search for global standards*. Business Horizons. https://www.sciencedirect.com/science/article/abs/pii/S0007681318301095

Ness, R. O. (2019, May 12). *Gödel's incompleteness theorems and the implications to building strong AI.* Towards Data Science. https://towardsdatascience.com/gödels-incompleteness-theorems-and-the-implications-to-building-strong-ai-1020506f6234

Netflix. (2021). *How Netflix's recommendations system works*. https://help.netflix.com/en/node/100639

Newsome, M. (2020, June 3). *Teletherapy in the Age of COVID-19*. North Carolina Health News. https://www.northcarolinahealthnews.org/2020/06/03/teletherapy-in-the-age-of-covid-19/

Norris, P. (2000). The worldwide digital divide. In *Paper for the Annaul Meeting of the Political Studies Association of the UK*. London School of Economics and Political Science.

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge University Press. doi:10.1017/CBO9781139164887

Novum Insights. (2021, June 30). *Macro overview of the crypto economy 2021 | Upcoming Novum Insights DeFi report 2021*. https://novuminsights.com/post/macro-overview-of-the-crypto-economy-2021-or-upcoming-novum-insights-defi-report-2021/

Nuñez, D., Agudo, I., & Lopez, J. (2017). Proxy Re-Encryption: Analysis of constructions and its application to secure access delegation. *Journal of Network and Computer Applications*, *87*, 193–209. doi:10.1016/j.jnca.2017.03.005

O'Connor, S. (2016, September 8). When your boss is an algorithm. *Financial Times*. https://www.ft.com/content/88fdc58e-754f-11e6-b60a-de4532d5ea35

O'Neil, C. (2016). Weapons of Math Destruction. Penguin Random House LLC.

O'Regan, C. (2018). *Hate speech regulation on social media: an intractable contemporary challenge*. Research Outreach. https://researchoutreach.org/wp-content/uploads/2020/02/Catherine-O-Regan.pdf

O'Rourke, B., Oortwijn, W., & Schuller, T.International Joint Task Group. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, *36*(3), 187–190. doi:10.1017/S0266462320000215 PMID:32398176

Office for Human Research Protections. (2016). *Informed consent FAQs*. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/faq/informed-consent/index.html

Ojimba, A. C., & Ikuli, B. Y. (2019). Friedrich Nietzsche's superman and its religious implications. *Journal of Philosophy*, *Culture and Religion*, *45*, 17–25. doi:10.7176/JPCR/45-03

Okaformbah, C. (2019, February 19). *Governance in a Decentralized Autonomous Organization*. Medium. https://just-charles.medium.com/governance-in-a-decentralized-autonomous-organization-425f56b3e8bb

Oppy, G., & Dowe, D. (2020). The turing test. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020 ed.). Stanford University. https://plato.stanford.edu/archives/win2020/entries/turing-test/

Oracle. (2019). *From fear to enthusiasm: Artificial intelligence is winning more hearts and minds in the workplace oracle and future workplace*. https://www.oracle.com/webfolder/s/assets/ebook/ai-work/conclusion.html#section5A

Orcutt, M. (2019, October 10). Why it's time to start talking about Blockchain ethics. *MIT Technology Review*. https://www.technologyreview.com/2019/10/10/132652/why-its-time-to-start-talking-about-Blockchain-ethics/

Orwat, C. (2019). *Risks of discrimination through the use of algorithms*. Federal Anti-Discrimination Agency. https://www.antidiskriminierungsstelle.de/EN/homepage/_documents/download_diskr_risiken_verwendung_von_algorithmen.pdf?__blob=publicationFile&v=1

Orwell, G. (1949). *1984*. Secker and Warburg.

Osman, H. (2019). *SA Health to overhaul EPAS. Healthcare IT news*. https://www.healthcareit.com.au/

Otterbach, S., Sousa-Poza, A., & Zhang, X. (2021). *Gender differences in perceived workplace harassment and gender egalitarianism: A comparative cross-national analysis*. Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1111/beer.12338

Oudshoorn, N. (2012). How Places Matter: Telecare Technologies and the Changing Spatial Dimensions of Healthcare. *Social Studies of Science*, *42*(1), 121–142. doi:10.1177/0306312711431817 PMID:22530385

Overby, S. (2020, May 7). *5 artificial intelligence (AI) types, defined*. The Enterprisers Project. https://enterprisersproject.com/article/2020/5/5-artificial-intelligence-ai-types-defined

Palmer, A. (2019, December 11). *Twitter CEO Jack Dorsey has an idealistic vision for the future of social media and is funding a small team to chase it*. CNBC. https://www.cnbc.com/2019/12/11/twitter-ceo-jack-dorsey-announces-bluesky-social-media-standards-push.html

Palmiter Bajorek, J. (2019). Voice Recognition still has significant race and gender biases. *Harvard Business Review*. Available at https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases

Pan American Health Organization World Health Organization. (2012). *Health Technology Assessment and Incorporation into Health Systems*. RESOLUTION CSP28.R9 Sixth meeting, 19 September. https://iris.paho.org/handle/10665.2/3684

Pareek, G., & Purushothama, B. R. (2020). Proxy re-encryption for fine-grained access control: Its applicability, security under stronger notions and performance. *Journal of Information Security and Applications*, *54*, 102543. doi:10.1016/j.jisa.2020.102543

Parmar, P., & Bhavsar, M. (2020). Achieving Trust using RoT in IaaS Cloud. *Procedia Computer Science*, *167*, 487–495. doi:10.1016/j.procs.2020.03.264

Parsons, S., & Abbott, C. (2013). Digital technologies for supporting the informed consent of children and young people in research: the potential for transforming current research ethics practice. *EPSRC Observatory for Responsible Innovation in ICT*. https://eprints.soton.ac.uk/id/eprint/356041

Parsons, S. (2015). The potential of digital technologies for transforming informed consent practices with children and young people in social research. *Social Inclusion (Lisboa)*, *3*(6), 56–68. doi:10.17645i.v3i6.400

Pascoe, C. (2011). Resource and Risk: Youth Sexuality and New Media Use. *Sexuality Research & Social Policy*, *8*(1), 5–17. doi:10.100713178-011-0042-5

Patel, V. (2012). Global Mental Health: From Science to Action. *Harvard Review of Psychiatry*, *20*(1), 6–12. doi:10.3109/10673229.2012.649108 PMID:22335178

Patranabis, S., Shrivastava, Y., & Mukhopadhyay, D. (2017). Provably Secure Key-Aggregate Cryptosystems with Broadcast Aggregate Keys for Online Data Sharing on the Cloud. *IEEE Transactions on Computers*, *66*(5), 891–904. doi:10.1109/TC.2016.2629510

Paul, K. (2019, October 25). The healthcare algorithm used across America has dramatic racial biases. *The Guardian*. https://www.theguardian.com/society/2019/oct/25/healthcare-algorithm-racial-biases-optum

Pecorino, P. A., & Maner, W. (1985). A proposal for a course on computer ethics. *Metaphilosophy*, *16*(4), 327–337. doi:10.1111/j.1467-9973.1985.tb00179.x

Peikoff, L. (2005). *Objectivism: The philosophy of Ayn Rand*. Penguin.

Perrin, A. (2019). Retrieved May 10, 2021 from Digital Gap between Rural and Nonrural America Persists. https://www.pewresearch.org/fact-tank/2019/05/31/digital-gap-between-rural-and-nonrural-america-persists/

Perry, W. L. (2013). *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. RAND Corporation. doi:10.7249/RR233

Pew Research Center. (2021, April 7). *Demographics of mobile device ownership and adoption in the United States*. https://www.pewresearch.org/internet/fact-sheet/mobile/

Philips-Wren, G., & Jain, L. (2006). Artificial intelligence for decision making. In B. Gabrys, R. J. Howlett, & L. C. Jain (Eds.), *Knowledge-based intelligent information and engineering systems* (pp. 531–536). Springer. doi:10.1007/11893004_69

Pieper, A. (1991). *Einführung in die Ethik*. UTB.

Polzer, J. T. (2018). Case study: Should an algorithm tell you who to promote. *Harvard Business Review*. https://hbr.org/2018/05/case-study-should-an-algorithm-tell-you-who-to-promote

Poole, G. A. (1996, Jan. 29). A new gulf in American education, the digital divide. *New York Times*, p. 2.

Popoveniuc, B. (2019). AIRSE: The ethics of artificially intelligent robots and systems. In A. Sandu, A. Frunza, & E. Unguru (Eds.), *Ethics in research practice and innovation* (pp. 283–295). IGI Global. doi:10.4018/978-1-5225-6310-5.ch015

Porter, G., Hampshire, K., Abane, A., Muthali, A., Robson, E., De Lannoy A., Tanle, A. & Owusu, A. (2020). Mobile Phones, Gender, and Female Empowerment in Sub-Saharan Africa: Studies with African Youth. *Information Technology for Development, 2*6(1), 180-93.

Posadas, D. V., Jr. (2018). The internet of things: The GDPR and the Blockchain may be incompatible. *Journal of Internet Law, 21*(11), 1, 20–29.

Potter, A. B. (2006). Zones of silence: A framework beyond the digital divide. *First Monday*, *11*(5). Advance online publication. doi:10.5210/fm.v11i5.1327

PricewaterhouseCoopers. (2020). *Internet advertising revenue report*. https://www.iab.com/wp-content/uploads/2020/05/FY19-IAB-Internet-Ad-Revenue-Report_Final.pdf

Prieger, J., & Hu, W. (2006). An empirical analysis of indirect network effects in the home video game market. Retrieved May 10, 2021, from: http://ssrn.com/abstract=941223 doi:10.2139/ssrn.941223

Publications Office of the European Union. (2020). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence act) and amending certain union legislative acts* (COM/2021/206 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

Puig de la Bellacasa, M. (2017). *Matters of Care: Speculative Ethics in More than Human Worlds*. University of Minnesota Press.

Puliafito, C., Mingozzi, E., Longo, F., Puliafito, A., & Rana, O. (2019). Fog Computing for the Internet of Things: A Survey. *ACM Trans. Internet Technol., 19*(2), 18:1-18:41. doi:10.1145/3301443

Pundir, N., Lindroos, M., McDonnell, J., Byrom, B., & Egan, S. (2020). Delving into econsent: Industry survey reinforces patient centricity. *Clinical Researcher (Alexandria, Va.)*, *34*(1). https://acrpnet.org/2020/01/14/delving-into-econsent-industry-survey-reinforces-patient-centricity/

Rabinow, P. (2008). Artificiality and Enlightenment: From Sociobiology to Biosociality. In J. X. Inda (Ed.), *Anthropologies of Modernity: Foucault, Governmentality, and Life Politics* (pp. 179–193). Blackwell Publishing.

Rahm, L. (2018). The ironies of digital citizenship: Educational imaginaries and digital losers across three decades. *Digital Culture & Society*, *4*(2), 39–61. doi:10.14361/dcs-2018-0204

Rand, A. (1963). *For the new intellectual*. Penguin.

Rand, A. (1964). *The virtue of selfishness*. Penguin.

Robeznieks, A. (2020, March 19). *Key changes made to telehealth guidelines to boost COVID-19 care*. American Medical Association. https://www.ama-assn.org/practice-management/digital/key-changes-made-telehealth-guidelines-boost-covid-19-care

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, *10*(1), 3069. Advance online publication. doi:10.103841467-019-10933-3 PMID:31337762

Rodriguez-Patarroyo, M., Torres-Quintero, A., Vecino-Ortiz, A. I., Hallez, K., Franco-Rodriguez, A. N., Rueda Barrera, E. A., Puerto, S., Gibson, D. G., Labrique, A., Pariyo, G. W., & Ali, J. (2020). Informed consent for mobile phone health surveys in Colombia: A qualitative study. *Journal of Empirical Research on Human Research Ethics*. doi:10.1177/1556264620958606 PMID:32975157

Roose, K. (2019, June 23). A machine may not take your job, but one could become your boss. *The New York Times*. https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html

Rosenbaum, E. (2019, April 3). *IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs*. CNBC. https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html

Rosenblat, A. (2018). *Uberland: How algorithms are rewriting the rules of work*. University of California Press. doi:10.1525/9780520970632

Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication*, *10*, 3758–3784. https://ijoc.org/index.php/ijoc/article/view/4892/1739

Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, *20*(2), 127–142. doi:10.100710676-018-9452-x

Rugeviciute, A., & Mehrpouya, A. (2019). Blockchain, a panacea for development accountability? A study of the barriers and enablers for Blockchain's adoption by development aid organizations. *Frontiers in Blockchain*, *2*, 15. Advance online publication. doi:10.3389/fbloc.2019.00015

Sachan, A., Emmanuel, S., & Kankanhalli, M. S. (2012). Aggregate Licenses Validation for Digital Rights Violation Detection. ACM Trans. *Multimedia Comput. Commun. Appl., 8*(2S), 37:1-37:21. doi:10.1145/2344436.2344443

Salamat, M. R. (2016). *Ethics of sustainable development: The moral imperative for the effective implementation of the 2030 agenda for sustainable development*. Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1111/1477-8947.12096

Salman, O., Elhajj, I., Chehab, A., & Kayssi, A. (2018). IoT survey: An SDN and fog computing perspective. *Computer Networks*, *143*, 221–246. doi:10.1016/j.comnet.2018.07.020

Sanderson, W. C., Arunagiri, V., Funk, A. P., Ginsburg, K. L., Krychiw, J. K., Limowski, A. R., Olesnycky, O. S., & Stout, Z. (2020). The Nature and Treatment of Pandemic-Related Psychological Distress. *Journal of Contemporary Psychotherapy*, *50*(4), 251–263. doi:10.100710879-020-09463-7 PMID:32836377

Sandler, R., & Basl, J. (2019). *Building data and AI ethics committees*. Accenture.

Sandvig, C., Hamilton, K., Karahalios, K., & Longbort, C. (2016). When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication*, *10*, 4972–4990. https://ijoc.org/index.php/ijoc/article/view/6182/1807

Satell, G. (2016, November 16). Teaching an algorithm to understand right and wrong. *Harvard Business Review*. https://hbr.org/2016/11/teaching-an-algorithm-to-understand-right-and-wrong

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., Eade, R., Fereday, D., Folland, C. K., Gordon, M., Hermanson, L., Knight, J. R., Lea, D. J., MacLachlan, C., Maidens, A., Martin, M., Peterson, A. K., Smith, D., Vellinga, M., ... Williams, A. (2014). Skillful Long-Range Prediction of European and North American Winters. Met Office Hadley Center. *Geophysical Research Letters*, *41*(7), 2514–2519. doi:10.1002/2014GL059637

Schmidt, F. A. (2017). Digital labour markets in the platform economy: *Mapping the political challenges of crowd work and gig work*. Friedrich-Ebert-Stiftung. https://library.fes.de/pdf-files/wiso/13164.pdf

Schmietow, B. (2016). Ethical dimensions of dynamic consent in data-intense biomedical research—Paradigm shift, or red herring? In D. Strech & M. Mertz (Eds.), *Ethics and Governance of Biomedical Research* (Vol. 4, pp. 197–209). Springer International Publishing. doi:10.1007/978-3-319-28731-7_15

Scott, B. (2018). Hard coding ethics into Fintech. *Finance & the Common Good / Bien Commun*, 44–45, 80–93.

Scribani, J. (2019, January 16). *What is Extended Reality (XR)?* Visual Capitalist. https://www.visualcapitalist.com/extended-reality-xr/

Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, *6*(3), 341–362. doi:10.1177/1461444804042519

Serrao, C., Marques, J., Dias, M., & Delgado, J. (2018). *Open-source software as a driver for digital content e-commerce and DRM interoperability*. Academic Press.

Setó-Pamies, D., & Papaoikonomou, E. (2020). *Sustainable development goals: A powerful framework for embedding ethics, CSR, and Sustainability in Management Education*. MDPI. https://www.mdpi.com/2071-1050/12/5/1762

Shah, M. (2019, August 21). *How Augmented Reality (AR) is changing the travel & tourism industry*. Towards Data Science. https://towardsdatascience.com/how-augmented-reality-ar-is-changing-the-travel-tourism-industry-239931f3120c

Shen, W., Yu, J., Xia, H., Zhang, H., Lu, X., & Hao, R. (2017). Light-weight and privacy-preserving secure cloud auditing scheme for group users via the third party medium. *Journal of Network and Computer Applications*, *82*, 56–64. doi:10.1016/j.jnca.2017.01.015

Shermer, M. (2004). *The science of good and evil: Why people cheat, gossip, care, share, and follow the golden rule*. Times Books.

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management*, *31*(2), 74–87. doi:10.4018/JDM.2020040105

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, *76*, 146–164. doi:10.1016/j.comnet.2014.11.008

Siero, N. B. (2017). *Guidelines for supporting teachers in teaching digital literacy* [Master's thesis, University of Twente]. University of Twente Theses Repository. https://essay.utwente.nl/73163/1/Siero_MA%20Educational%20Science%20And%20Technology_BMS.pdf

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. doi:10.1038/nature16961 PMID:26819042

Simon, C. M., Schartz, H. A., Rosenthal, G. E., Eisenstein, E. L., & Klein, D. W. (2018). Perspectives on electronic informed consent from patients underrepresented in research in the United States: A focus group study. *Journal of Empirical Research on Human Research Ethics; JERHRE*, *13*(4), 338–348. doi:10.1177/1556264618773883 PMID:29790410

Simon, J. (2017). Value-sensitive design and responsible research and innovation. In S. O. Hansson (Ed.), *The ethics of technology* (pp. 219–236). Rowman & Littlefield International.

Simpson, L., Daws, L., & Pini, B. (2004). Public internet access revisited. *Telecommunications Policy*, *28*(3), 323–337. doi:10.1016/j.telpol.2003.10.001

Sindwani, P. (2020, January 21). *The boss machine is here — AI is all set to eliminate middle management in 8 years*. Business Insider. https://www.businessinsider.in/careers/news/the-boss-machine-is-here-ai-is-all-set-to-eliminate-middle-managers-in-8-years/articleshow/73474729.cms

Singer, P. (1975). *Animal liberation*. Random House.

Singh, A. K., Nag, A., Karforma, S., & Mukhopadhyay, S. (2019). Implementation of multi-agent based Digital Rights Management System for Distance Education (DRMSDE) using JADE. *International Journal of Advanced Computer Science and Applications*, *10*(3), 343–352. doi:10.14569/IJACSA.2019.0100345

Sirico, R. A. (1997). *The Religious Left with Its Mask Off*. https://www.acton.org/

Sirk, C. (2020, August 21). *Diffusion of innovation: How adoption of new tech spreads*. CRM. https://crm.org/articles/diffusion-of-innovations

Sisto, D. (2021). *Remember me: Memory and forgetting in the digital age* (A. Killgarriff, Trans.). Wiley.

Skelton, E., Drey, N., Rutherford, M., Ayers, S., & Malamateniou, C. (2020). Electronic consenting for conducting research remotely: A review of current practice and key recommendations for using e-consenting. *International Journal of Medical Informatics*, *143*, 104271. doi:10.1016/j.ijmedinf.2020.104271 PMID:32979650

Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., Watson, Z., Breen, G., Schwarz, R., Steptoe, W., Szostak, D., Halan, S., Fox, D., & Silver, J. (2020). The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality*, *1*, 1. Advance online publication. doi:10.3389/frvir.2020.00001

Smith, T. (2006). *Ayn Rand's normative ethics: The virtuous egoist*. Cambridge University Press. doi:10.1017/CBO9781139167352

Solsman, J. E. (2018, January 10). *YouTube's AI is the puppet master over most of what you watch*. CNet. https://www.cnet.com/news/youtube-ces-2018-neal-mohan/

Spacey, J. (2016). *20 Types of Technology Ethics*. Simplicable. https://simplicable.com/

Sparrow, R. (2016). Robots and respect: Assessing the case against autonomous weapon systems. *Ethics & International Affairs*, *30*(1), 93–116. doi:10.1017/S0892679415000647

Sparviero, S., & Ragnedda, M. (2021). Towards digital sustainability: the long journey to the sustainable development goals 2030. *Digital Policy, Regulation and Governance*.

Spruit, S. L., van de Poel, I., & Doorn, N. (2016). Informed consent in asymmetrical relationships: An investigation into relational factors that influence room for reflection. *NanoEthics*, *10*(2), 123–138. doi:10.100711569-016-0262-5 PMID:27478516

Statista. (2021). Retrieved May 10, 2021 from www.statista.com/

Stroponiati, K., Abugov, I., Varelas, Y., Stroponiatis, K., Jurgeleviciene, M., & Rao, Y. S. R. (2020). Decentralized governance in DeFi: Examples and pitfalls. *Squarespace.* https://static1.squarespace.com/static/5966eb2ff7e0ab3d29b6b55d/t/5f989987fc086a1d8482ae70/1603837124500/defi_governance_paper.pdf

Subramanyam, A. V., Emmanuel, S., & Kankanhalli, M. S. (2012). Robust Watermarking of Compressed and Encrypted JPEG2000 Images. *IEEE Transactions on Multimedia*, *14*(3), 703–716. doi:10.1109/TMM.2011.2181342

Suciu, P. (2019, October 11). More Americans are getting their news from social media. *Forbes*. https://www.forbes.com/sites/petersuciu/2019/10/11/more-americans-are-getting-their-news-from-social-media/?sh=578f1ca43e17

Sulkowski, A. (2020). *Tao OF Dao: Hardcoding business ethics on blockchain*. https://www.researchgate.net/profile/Adam-Sulkowski-2/publication/336316096_THE_TAO_OF_DAO_HARDCODING_BUSINESS_ETHICS_ON_BLOCKCHAIN/links/5e8dd07ba6fdcca789fe0a34/THE-TAO-OF-DAO-HARDCODING-BUSINESS-ETHICS-ON-BLOCKCHAIN.pdf

Sultan, N. H., Barbhuiya, F. A., & Laurent, M. (2018). ICAuth: A secure and scalable owner delegated inter-cloud authorization. *Future Generation Computer Systems*, *88*, 319–332. doi:10.1016/j.future.2018.05.066

Sumpter, D. (2018). *Outnumbered: From Facebook and Google to fake news and filter-bubbles – The algorithms that control our lives*. Bloomsbury Sigma.

Sun, L., Xu, C., Li, C., & Li, Y. (2020). Server-aided searchable encryption in multi-user setting. *Computer Communications*, *164*, 25–30. doi:10.1016/j.comcom.2020.09.018

Surthi, S. (2020). *New paradigms in business management practices*. https://www.amazon.in/NEW-PARADIGMS-BUSINESS-MANAGEMENT-PRACTICES-ebook/dp/B08RQR5Q9N/ref=sr_1_2?dchild=1&qid=1633022006&qsid=260-5567715-1581557&refinements=p_27%3ASruthi%2BS&s=digital-text&sr=1-2&sres=B07KKKK7NQ%2CB08RQR5Q9N%2CB08RP1MX5P%2CB08R7XQRFW%2CB07DPMZ1PZ&text=Sruthi%2BS

Suryansh, S. (2018, March 26). *Genetic algorithms + neural networks = Best of both worlds.* Torwards Data Science. https://towardsdatascience.com/gas-and-nns-6a41f1e8146d

Tait, A. R., & Voepel-Lewis, T. (2015). Digital multimedia: A new approach for informed consent? *Journal of the American Medical Association*, *313*(5), 463–464. doi:10.1001/jama.2014.17122 PMID:25647199

Takahashi, D. (2020, October 27). *Alethea AI makes it easy to create AI avatars from a single photo*. VentureBeat. https://venturebeat.com/2020/10/27/alethea-ai-makes-it-easy-to-create-ai-avatars-from-a-single-photo/

Tang, Y., Xiong, J., Becerril-Arreola, R., & Iyer, L. (2019). Ethics of Blockchain: A framework of technology, applications, impacts, and research directions. *Information Technology & People*, *33*(2), 602–632. doi:10.1108/ITP-10-2018-0491

Tapas, N., Longo, F., Merlino, G., & Puliafito, A. (2020). Experimenting with smart contracts for access control and delegation in IoT. *Future Generation Computer Systems*, *111*, 324–338. doi:10.1016/j.future.2020.04.020

Taylor, R. (2020). *Is the algorithm working for us? Algorithms, qualifications and fairness*. Centre for Progressive Policy. https://www.progressive-policy.net/downloads/files/Is-the-algorithm-working-for-us_Roger-Taylor.pdf

Thanh, T. M., & Iwakiri, M. (2016). Fragile watermarking with permutation code for content-leakage in digital rights management system. *Multimedia Systems*, *22*(5), 603–615. doi:10.100700530-015-0472-7

The American Medical Association (AMA). (2018). *AMA adopts a policy on augmented intelligence*. https://medicalxpress.com/news

The Canadian Agency for Drugs and Technologies in Health CADTH. (2015). *Health Technology Expert Review Panel: Process for Developing Recommendations.* Retrieved from https://www.cadth.ca/sites/default/files/pdf/HTERP_Process.pdf

The Economist. (2018, June 21). *How an algorithm may decide your career*. https://www.economist.com/business/2018/06/21/how-an-algorithm-may-decide-your-career

The Economist. (2019, October 5). *The stock market is now run by computers, algorithms, and passive managers*. https://www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers

The European Coordination Committee of the Radiological, Electromedical and Healthcare IT Industry (COCIR). (2019). *European Health Data Space: Towards a Better Patient Outcome*. https://www.cocir.org/fileadmin/Publications

The European Medicines Agency's (EMA). (2020). *Regulatory science strategy*. https://www.ema.europa.eu/

The International Alliance of Patients' Organizations (IAPO). (2017). https://www.iapo.org.uk

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research* ((OS) 78-0013). U.S. Department of Health, Education, and Welfare. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf

The Pharmaceutical Group of the European Union (PGEU). (2016). *Position P aper on Big Data & Artificial Intelligence in Healthcare.* Ref 19.02.20E 001. https://www.pgeu.eu/

Thierer, A. (2019). *Global Innovation Arbitrage: Export Controls Edition*. The Technology Liberation Front. https://techliberation.com/

Third, A., Collin, P., Walsh, L., & Black, R. (2019). *Young People in Digital Society: Control Shift*. Palgrave MacMillan. doi:10.1057/978-1-137-57369-8

Thomason, J. (2021, May 23). *Why DAO Governance Matters for DeFi* [Article]. LinkedIn. https://www.linkedin.com/pulse/why-dao-governance-matters-defi-dr-jane-thomason

Thomason, J., Bernhardt, S., Kansara, T., & Cooper, N. (2019). *Blockchain technology for global social change*. IGI Global. doi:10.4018/978-1-5225-9578-6

Thornill, J. (2019). Should we think of Big Tech as Big Brother? *Financial Times*. Available at https://www.ft.com/content/43980f9c-0f5b-11e9-a3aa-118c761d2745

Time Doctor. (2020). *Time tracking software to help your team be more productive while working from home*. https://www.timedoctor.com/

Todt, G., Weiss, M., & Hoegl, M. (2018). Mitigating negative side effects of innovation project terminations: The role of resilience and social support. *Journal of Product Innovation Management*, *35*(4), 518–542. doi:10.1111/jpim.12426

Tolmeijer, S., Kneer, M., Sarasua, C., Christen, M., & Bernstein, A. (2020). Implementations in machine ethics: A survey. *ACM Computing Surveys*, *53*(6), 132. Advance online publication. doi:10.1145/3419633

Torres, V., Serrao, C., Dias, M. S., & Delgado, J. (2008). Open DRM and the future of media. *IEEE MultiMedia*, *15*(2), 28–36. doi:10.1109/MMUL.2008.38

Trevelli, C., & Morel, J. (2019). *Rural Youth Inclusion, Empowerment and Participation* (Report No. 45). International Fund for Agricultural Development. www.ifad.org/ruraldevelopmentreport

Tronto, J. (1993). *Moral Boundaries: A Political Argument for an Ethic of Care*. Routledge.

Tronto, J. C. (2013). *Caring Democracy: Markets, Equality, and Justice*. NYU Press.

Tsai, H. J., & Wu, Y. (2021). Changes in Corporate Social Responsibility and Stock Performance. *Journal of Business Ethics*. Advance online publication. doi:10.100710551-021-04772-w

Turner, A. (2021, August). *How many smartphones are in the world?* BankMyCell. https://www.bankmycell.com/blog/how-many-phones-are-in-the-world

Tutty, S., Spangler, D. L., Poppleton, L. E., Ludman, E. J., & Simon, G. E. (2010). Evaluating the effectiveness of cognitive-behavioral teletherapy in depressed adults. *Behavior Therapy*, *41*(2), 229–236. doi:10.1016/j.beth.2009.03.002 PMID:20412887

Tzanou, M. (2020). *Health data privacy under the GDPR: Big data challenges and regulatory responses*. Routledge. doi:10.4324/9780429022241

Uniform Law Commission. (1999). *Uniform Electronic Transaction Act*. http://euro.ecom.cmu.edu/program/law/08-732/Transactions/ueta.pdf

United Nations Convention on the Rights of the Child. (1989).

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2014). *Reading in the Mobile Era: A Study of Mobile Reading in Developing Countries.* https://unesdoc.unesco.org/ark:/48223/pf0000227436

United Nations Human Rights Office of the High Commissioner. (n.d.). https://www.ohchr.org/EN/Issues/Pages/WhatareHumanRights.aspx

United Nations Sustainable Development Goals. (n.d.). Retrieved May 10, 2021 from https://sdgs.un.org/goals/goal17

United Nations. (n.d.). *Millennium development goals*. https://www.un.org/millenniumgoals/

United States Department of Commerce's National Telecommunications and Information Administration. (n.d.). *Internet and Computer Use Studies and Data Files*. Retrieved May 10, 2021 from https://www.ntia.doc.gov/data

University of Illinois. (2020). *Emerging Technologies and Their Impact on Health Informatics*. https://healthinformatics.uic.edu/

Vallor, S. (2018). *Ethical toolkit*. https://www.scu.edu/ethics-in-technology-practice/ethical-toolkit/

Van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, *19*(4), 315–326. doi:10.1080/01972240309487

van Es, K. (2020). You Tube's operational logic: "The view" as pervasive category. *Television & New Media*, *21*(3), 223–239. doi:10.1177/1527476418818986

van Hooijdonk, R. (2019, November 27). *AI in management: Your boss could soon be a machine*. https://blog.richardvanhooijdonk.com/en/ai-in-management-your-boss-could-soon-be-a-machine/

Van Leeuwen, D. (2020). *Bioethics: Autonomy. For me, on behalf of me: An interview with Kenneth Goodman*. Health Hats. https://www.health-hats.com/wp-content/uploads/2020/02/20200209_HHP059_Kenneth_Goodman_FINAL.pdf

van Rijmenam, M. (2020). *Algorithmic management: What is it (and what's next)?* Medium. https://medium.com/swlh/algorithmic-management-what-is-it-and-whats-next-33ad3429330b

Vandaele, K. (2018). *Will trade unions survive in the platform economy? Emerging patterns of platform workers' collective voice and representation in Europe* (Working Paper No. 2018.05). European Trade Union Institute. https://www.etui.org/publications/working-papers/will-trade-unions-survive-in-the-platform-economy-emerging-patterns-of-platform-workers-collective-voice-and-representation-in-europe

Varadan, S. (2019). The Principle of Evolving Capacities under the UN Convention on the Rights of the Child. *International Journal of Children's Rights*, *27*(2), 306–338. doi:10.1163/15718182-02702006

Varkey, B. (2021). Principles of clinical ethics and their application to practice. *Medical Principles and Practice*, *30*(1), 17–28. PMID:32498071

Vayena, E., & Blasimme, A. (2017). Biomedical big data: New models of control over access, use and governance. *Journal of Bioethical Inquiry*, *14*(4), 501–513. doi:10.100711673-017-9809-6 PMID:28983835

Veatch, R. (2008). *Patient, Heal Thyself: How the "New Medicine" Puts the Patient in Charge*. Oxford University Press.

Verhellen, E. (2015). The Convention on the Rights of the Child: Reflections from a Historical, Social Policy and Educational Perspective. In W. Vanderhole, E. Desmet, D. Raynaert, & S. Lambrechts (Eds.), *Routledge International Handbook of Children's Rights Studies* (pp. 43–59). Routledge.

Verma, G. K., & Singh, B. B. (2018). Efficient identity-based blind message recovery signature scheme from pairings. *IET Information Security*, *12*(2), 150–156. doi:10.1049/iet-ifs.2017.0342

Vezyridis, P., & Timmons, S. (2019). Resisting big data exploitations in public healthcare: Free riding or distributive justice? *Sociology of Health & Illness*, *41*(8), 1585–1599. doi:10.1111/1467-9566.12969 PMID:31423602

Victorian Law Reform Commission. (2008). *Law of abortion: Final report. Appendix B – Ethics of abortion.* https://www.lawreform.vic.gov.au/wp-content/uploads/2021/07/VLRC_Abortion_Report-1.pdf

Vidal, E. (2018, April 9). *An explosion in people analytics*. Converge. https://convergetechmedia.com/an-explosion-in-people-analytics/

Vidgen, R., Hindle, G., & Randolph, I. (2020). Exploring the ethical implications of business analytics with a business ethics canvas. *European Journal of Operational Research*, *281*(3), 491–501. doi:10.1016/j.ejor.2019.04.036

Vieth, A. (2006). *Einführung in die Angewandte Ethik*. WGB.

Vilardell, F. (1990). Ethical problems of medical technology. *Bulletin of the Pan American Health Organization*, *24*(4), 379–385. PMID:2073552

Voshmgir, S. (2019, July). Tokenized networks: What is a DAO? *BlockchainHub*. https://blockchainhub.net/dao-decentralized-autonomous-organization/

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. doi:10.1126cience.aap9559 PMID:29590045

Voundi Koe, A. S., & Lin, Y. (2019). Offline privacy preserving proxy re-encryption in mobile cloud computing. *Pervasive and Mobile Computing*, *59*, 101081. doi:10.1016/j.pmcj.2019.101081

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. Advance online publication. doi:10.1016/j.clsr.2021.105567

Wall, J. (2017). *Children's Rights: Today's Global Challenge*. Rowman & Littlefield.

Walsh, M. (2019, May 8). When algorithms make managers worse. *Harvard Business Review*. https://hbr.org/2019/05/when-algorithms-make-managers-worse

Wamba-Taguimdje, S.-L., Fosso Wamba, S., Kala Kamdjoug, J. R., & Tchatchouang Wanko, C. E. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal*, *26*(7), 1893–1924. doi:10.1108/BPMJ-10-2019-0411

Ward, L. (2015). Caring for Ourselves? Self-Care and Neoliberalism. In M. Barnes, T. Brannelly, L. Ward, & N. Ward (Eds.), *Ethics of Care: Critical Advances in International Perspective*. Policy Press. doi:10.2307/j.ctt1t89d95.8

Warschauer, M. (2002). Reconceptualizing the digital divide. *First Monday*, *7*(7). Advance online publication. doi:10.5210/fm.v7i7.967

Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT press. doi:10.7551/mitpress/6699.001.0001

Warschauer, M. (2012). The digital divide and social inclusion. *American Quarterly*, *6*(2), 131.

Watkins, C. (2018a). Preface. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Divide: How Black and Latino Youth Navigate Digital Inequality* (pp. ix–xiii). NYU Press.

Watkins, C. (2018b). How Black and Latino Youth are Remaking the Digital Divide. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Edge: How Black and Latino Youth Navigate Digital Inequality* (pp. 19–49). NYU Press.

Watkins, C. (2018c). The Mobile Paradox: Understanding the Mobile Lives of Latino and Black Youth. In C. Watkins, A. Lombana-Bermudez, A. Cho, J. Ryan Vickery, V. Shaw, & L. Weinzimmer (Eds.), *The Digital Edge: How Black and Latino Youth Navigate Digital Inequality* (pp. 50–77). NYU Press.

Weatherly, N. L. (2021). The ethics of organizational Behavior Management. *Journal of Organizational Behavior Management*, *41*(3), 197–214. doi:10.1080/01608061.2021.1890664

Wee, R., Henaghan, M., & Winship, I. (2013). Ethics: Dynamic consent in the digital age of biology: Online initiatives and regulatory considerations. *Journal of Primary Health Care*, *5*(4), 341–347. doi:10.1071/HC13341 PMID:24294625

West, D. M., & Allen, J. R. (2020). *Turning Point. Policymaking in the Era of Artificial Intelligence*. The Brookings Institution.

Whiting, R., & Pritchard, K. (2017, December 1). *Digital Ethics*. Birkbeck Institutional Research Online. https://eprints.bbk.ac.uk/id/eprint/15623/

Whitney, L. (2017, September 29). Are computers already smarter than humans? *Time Magazine*. http://time. com/4960778/computers-smarter-than-humans/

Whooley, O. (2017). Defining Mental Disorders: Sociological Investigations into the Classification of Mental Disorders. In T. L. Scheid & E. R. Wright (Eds.), *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems* (pp. 45–65). Cambridge University Press. doi:10.1017/9781316471289.006

Wiener, J. B. (2004). The regulation of technology, and the technology of regulation. *Technology in Society*, *26*(2-3), 483–500. doi:10.1016/j.techsoc.2004.01.033

Wilbanks, J. T. (2020). Electronic informed consent in mobile applications research. *The Journal of Law, Medicine & Ethics*, *48*(1, suppl), 147–153. doi:10.1177/1073110520917040 PMID:32342737

Wilson, J., & Wilson, H. (2009). Digital Divide: Impediment to ICT and Peace Building in Developing Countries. *American Communication Journal*, *11*(2), 1–9.

Win, L. L., Thomas, T., & Emmanuel, S. (2012). Privacy Enabled Digital Rights Management Without Trusted Third Party Assumption. *IEEE Trans. Multimedia, 14*(3–1), 546–554. doi:10.1109/TMM.2012.2189983

Winner, L. (1980). Do artifacts have politics? *Daedalus*, *109*(1), 121–136.

Wisskirchen, G., Biacabe, B. T., Bormann, U., Muntz, A., Niehaus, G., Soler, G. J., & von Brauchitsch, B. (2017). *Artificial intelligence and robotics and their impact on the workplace*. IBA Global Employment Institute.

Wolchover, N. (2020, July 14). How Gödel's proof works. *Quanta Magazine*. https://www.quantamagazine.org/how-godels-incompleteness-theorems-work-20200714/

Woo, B., Walton, E., & Takeuchi, D. T. (2017). Cultural Diversity and Mental Health Treatment. In T. L. Scheid & E. R. Wright (Eds.), *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems* (pp. 493–511). Cambridge University Press. doi:10.1017/9781316471289.029

Wood, A. J. (2021). *Algorithmic management: Consequences for work organisation and working conditions* (Working paper). Joint Research Centre, European Commission. https://ec.europa.eu/jrc/sites/default/files/jrc124874.pdf

World Bank. (2019). *World development report 2019: The changing nature of work*. https://www.worldbank.org/en/publication/wdr2019

World Economic Forum. (2019, March). *Central banks and distributed Ledger technology: How are central banks exploring Blockchain today?* [Whitepaper]. https://www3.weforum.org/docs/WEF_Central_Bank_Activity_in_Blockchain_DLT.pdf

World Health Organization (WHO). (2015). *Global Survey on Health Technology Assessment by National Authorities.* www.who.int/health-technology-assessment

World Health Organization (WHO). (2021). https://www.who.int/

Xenidis, R., & Senden, L. (2020). EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination. In General Principles of EU law and the EU Digital Order. Kluwer Law International.

Xie, M., Ruan, Y., Hong, H., & Shao, J. (2021). A CP-ABE scheme based on multi-authority in hybrid clouds for mobile devices. *Future Generation Computer Systems*, *121*, 114–122. doi:10.1016/j.future.2021.03.021

Xu, R. Z., Zhang, L., Zhao, H. W., & Peng, Y. (2017). Design of Network Media's Digital Rights Management Scheme Based on Blockchain Technology. *2017 IEEE 13th International Symposium on Autonomous Decentralized Systems (Isads 2017),* 128–133. 10.1109/ISADS.2017.21

Yang, J., & Basile, K. (2021, April 20). Communicating corporate social responsibility: External stakeholder involvement, productivity and firm performance. *Journal of Business Ethics*. https://link.springer.com/article/10.1007/s10551-021-04812-5

Yang, G. (2013). *The power of the Internet in China: Citizen activism online*. Columbia University Press.

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(4), 1036–1040. doi:10.1073/pnas.1418680112 PMID:25583507

Zafar, F., Khan, A., Suhail, S., Ahmed, I., Hameed, K., Khan, H. M., Jabeen, F., & Anjum, A. (2017). Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes. *Journal of Network and Computer Applications*, *94*, 50–68. doi:10.1016/j.jnca.2017.06.003

Zalatimo, S. (2021, April 13). *Coming summer 2021*. Voice. https://about.voice.com/blog/launching-summer-2021-nft-socialmedia-creators/

Zayyad, H. M. (n.d.). *Corporate Social Responsibility and patronage intentions: The mediating effect of Brand Credibility*. Taylor & Francis. https://www.tandfonline.com/doi/abs/10.1080/13527266.2020.1728565

Zelizer, V. A. (2009). *The Purchase of Intimacy*. Princeton University Press. doi:10.1515/9781400826759

Zhang, Z. Y., Wang, Z., & Niu, D. M. (2015). A novel approach to rights sharing-enabling digital rights management for mobile multimedia. *Multimedia Tools and Applications*, *74*(16), 6255–6271. doi:10.100711042-014-2135-7

Zhaofeng, M., Weihua, H., & Hongmin, G. (2018). A new blockchain-based trusted DRM scheme for built-in content protection. *Eurasip Journal on Image and Video Processing, 2018*(1). doi:10.1186/s13640-018-0327-1

Zhou, Y., Chen, C., & Johansson, M.J.O. (2013). The pre-mRNA retention and splicing complex controls tRNA maturation by promoting TAN1 expression. *Nucleic Acids Research, 41*(11), 5669–5678. doi:10.1093/nar/gkt269

Zhou, X., Snoswell, C. L., Harding, L. E., Bambling, M., Edirippulige, S., Bai, X., & Smith, A. C. (2020). The Role of Telehealth in Reducing the Mental Health Burden from COVID-19. *Telemedicine Journal and e-Health*, *26*(4), 377–379. doi:10.1089/tmj.2020.0068 PMID:32202977

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. Hachette Book Group.

Zwitter, A., & Hazenberg, J. (2020). Decentralized network governance: Blockchain technology and the future of regulation. *Frontiers in Blockchain*, *3*, 12. Advance online publication. doi:10.3389/fbloc.2020.00012

About the Contributors

Ingrid Vasiliu-Feltes is an Expert Advisor to the EU Blockchain Observatory Forum, serves as a Board Member of UN Legal and Economic Empowerment Network. Founder and CEO of The Science, Entrepreneurship and Investments Institute. Country Director for WBAF USA, Senator of WBAF, Faculty Member of the WBAF Business School-Division of Entrepreneurship, and teaching the Executive MBA Business Technology Course at the UM Business School. Chief Innovation Officer and past Chief Ethics Officer for GBA Global, Senator and as the Chief Quality and Safety Officer at MEDNAX and provides leadership to the MEDNAX Center for Research, Education, Quality and Safety. During her academic tenure she taught 50 + courses and served Chief of Compliance and Director of Research. Member of numerous prestigious professional organizations. She holds several certifications, such as Bioethics from Harvard, AI and Business Strategy from MIT Sloan, Blockchain Technology and Business Innovation from MIT Sloan, Finance from Harvard Business School, Negotiation from Harvard Law School, Innovation and Entrepreneurship from Stanford Graduate School of Business, Fellow of the American College of Healthcare Executives, Master Black Belt in Lean and Six Sigma Management.

Jane Thomason is a visionary leader and social innovator driving digital transformation to reduce inequality and shift to a human centred digital economy where: we measure value based on people and the planet; we measure and report on impact; we build a new data model with increased individual control and ability to monetise; and we eliminate social economic disparity. Successful CEO and Founder with a track record of leadership globally including: CEO of tertiary teaching hospital; Chair of the Board of leading private hospital; senior social sector manager Asian Development Bank and founder and CEO of successful international development company which reached $250 million revenue under her leadership. Non Executive Director ASX listed bank.

* * *

Anna Amsler is an independent consultant and researcher affiliated to the Observatory of Competitiveness and New Ways of Working. She holds a Bachelor's Degree in International Relations and a Master's in Political Communication and Marketing, having worked in private and public institutions in areas related to public policy, strategic planning and project evaluation.

Ivana Bartoletti is a privacy, data protection and digital ethics professional and thought leader. She is the Founder of the Women Leading in AI Network, Author of An Artificial Revolution: on Power, Politics and AI and a Visiting Policy Fellow at the Oxford Internet Institute, University of Oxford.

Wendy Charles has been involved in clinical trials from every perspective for 30 years with a strong background in operations and regulatory compliance. She currently serves as Chief Scientific Officer for BurstIQ, a healthcare-oriented blockchain company, where she leads the Life Sciences division. Wendy augments her blockchain healthcare experience by serving on the EU Blockchain Observatory and Forum Expert Panel, HIMSS Blockchain Task Force, Government Blockchain Association healthcare group, IEEE Blockchain working groups, and she is involved as an assistant editor and reviewer for academic journals. Wendy obtained her PhD in Clinical Science with a specialty in Health Information Technology from the University of Colorado, Anschutz Medical Campus. Wendy is certified as an IRB Professional, Clinical Research Professional, and Blockchain Professional.

Odilia Coi is a creative, unconventional thinker and problem-solver with a multi-disciplinary academic background. She holds a BA in Philosophy with a dissertation on moral implication in classic liberalism economic paradigm. As an extremely curious person, she also earned a B.Sc in Physics Eng. and a M.Sc in Electronics. During her Ph.D., she worked with French Space Agency to push the boundaries of knowledge and technology ensuring successful space missions. An IEEE member, she has been a speaker at various international conferences sponsored by NASA, and she has authored and co-authored several publications in prestigious Scientific Journals. In May 2021, she has joined MKAI leadership team as Programme Manager.

Robin Craig has a PhD in molecular genetics, and cofounded biotechnology company Genesearch, now one of Australia's oldest privately owned biotech firms. He has wide interests in science and philosophy and in addition to numerous scientific papers has written a long-running series on philosophy for the magazine of the high-IQ society Mensa in Australia and is a published author on ethics. He has given numerous talks on scientific topics with philosophical interest, including the ethics of artificial intelligence, to community and professional groups and on cruise ships.

David Danks is Professor of Data Science & Philosophy and affiliate faculty in Computer Science & Engineering at University of California, San Diego. His research interests are at the intersection of philosophy, cognitive science, and machine learning, using ideas, methods, and frameworks from each to advance our understanding of complex, interdisciplinary problems. Danks has examined the ethical, psychological, and policy issues around AI and robotics in transportation, healthcare, privacy, and security. He has also done significant research in computational cognitive science, culminating in his Unifying the Mind: Cognitive Representations as Graphical Models (2014, The MIT Press). And he has developed multiple novel causal discovery algorithms for complex types of observational and experimental data. Danks is the recipient of a James S. McDonnell Foundation Scholar Award, as well as an Andrew Carnegie Fellowship. He received an A.B. in Philosophy from Princeton University, and a Ph.D. in Philosophy from University of California, San Diego. He was previously the L.L. Thurstone Professor of Philosophy & Psychology at Carnegie Mellon University, as well as the Chief Ethicist of CMU's Block Center for Technology & Society; co-director of CMU's Center for Informed Democracy and Social Cybersecurity (IDeaS); and an adjunct member of the Heinz College of Information Systems and Public Policy.

Muhammet Demirbilek is an Associate Professor of Educational Technology in the Faculty of Education at Suleyman Demirel University in Isparta /Turkey. He earned his doctoral and master's degree in Educational Technology from the University of Florida. He also holds B.S. and M.S. degrees in electronics engineering from Istanbul University. He worked as a graduate faculty and visiting Assistant Professor in Digital Worlds Institute at University of Florida (2013-2014) and Post-Doctoral Researcher at Games, Learning, and Society (GLS) at the University of Wisconsin-Madison (2008-2009). His research interests include the impact of digital media and computer games and simulations on teaching and learning. How new media (e.g web 2.0, online social networks, mobile media) . His recent research interests are the effects of online bullying, learning analytic, big data analysis, using mobile games to improve social skills of kids with autism. How augmented reality games on handheld computers designed and used for formal and informal education. How social media, online social networks, and mobile media restructure thinking, values, actions, education, community, and culture.

Patrick Flanagan is Associate Professor and Chair of Theology and Religious Studies at St. John's University, Jamaica, NY. There Patrick teaches courses in the moral theology of healthcare and the marketplace. Patrick's area of research focuses on the intersection of religious ethics and information technology. For his doctoral degree from Loyola University, Chicago, IL, Patrick wrote his dissertation on the import of an ethic of the common good for a technological society. Since receipt of his doctorate, Patrick has presented at numerous conferences and written articles on union labor, digital rights, the role of the Catholic pharmacist, cybersex, and the digital divide. Presently he is working on a book on the seven deadly sins of information technology.

Richard Foster-Fletcher is an Artificial Intelligence (AI) Advisor, Author and Speaker. Drawing on his experiences of working in bluechip technology companies, he founded NeuralPath.io, an AI Consulting Practice that advises companies on how to responsibly monetize data. Richard is the Founder and Chair of the MKAI Network and Host of the Boundless Podcast. He speaks about AI innovation, ethics and the future of work, and regularly lectures on the subject of Responsible AI, notably delivering talks for LSE, UCL and Cranfield University. A graduate of the MIT AI Strategy Course and a Wiley published author on the Future of AI, Richard is a member of the Resilience Frontiers Advisory Group for Communicating Climate Change and has collaborated with the United Nations Environmental Programme (UNEP) to support their AI and Digital Transformation journey.

Ali Hussain is a postgraduate student at the University of Malaya. He focuses on research and development of next generation protocols and technology review and assessment. He studied at the National University of Science and Technology, Islamabad, Pakistan.

Ritesh Jain is an Entrepreneurial Technology Leader, Board Advisor, start-up founder and a PhD in Payments Innovation, Open Banking and Open Finance. Over the last 20 years, Ritesh has been at the forefront of the payments industry, leading the future of VISA payments and introducing Apple Pay in Europe / UK. He has led the transformation across many global organisations. Ritesh's vast experience and expertise in the payments industry is now seeing him play crucial advisory roles to many organisations, such as the G20 Initiative for Financial Inclusion (Member); Open Banking in EU and Africa; Harvard Business Review; payment regulators and government bodies for their initiative in social and financial inclusion; MIT Global in Tech Panel (Member). He has also been a mentor to the UK Parliament Digital initiative and many other start-ups. Ritesh is also an advisor and visiting lecturer to leading business schools and universities globally, apart from being a regular speaker and author and a strong advocate of diversity and inclusion. After leaving HSBC as COO Digital, Ritesh co-founded Infynit, harnessing his knowledge of and penchant for digital payments.

Lucia Lucchini is a Manager in the Cyber Risk team at Deloitte and specialises on AI governance and policy. She also sits on the steering committee of the global Women for Ethical AI Network at Deloitte.

Ruth Magtanong is a research manager for the Department of Medicine at Denver Health and Hospital Authority. She is simultaneously completing a doctoral degree at Case Western Reserve University. For more than 15 years, she has served in a variety of research roles navigating informed consent processes for international and US-based clinical research studies.

Monika Manolova holds a PhD in geographic information systems (GIS) from Sofia University 'St. Kliment Ohridski', is a cum laude graduated with excellence as M.Sc. in planning and governance of territorial systems and a bachelor in regional development and politics. Monika has experience as GIS expert on several complex regional projects. She has expertise with a multitude of specialized GIS software systems ArcGIS, QGIS, MapInfo, GeoServer, MapServer, PostGIS, OpenLayers. GDAL for raster data. Has multiple certifications in Microsoft Azure, Microsoft SQL Server, C#, Python coding and visualization for data science. Apart from her technical background, Monika is a Digital ecosystems expert for the Digital National Alliance in Bulgaria and a member of the DIH AgroHub BG with a focus on inclusion, equality, and smart growth for communities. Monika is a member of the EIP SCC group on smart mobility and EIP SCC group on citizen control of personal data, an MKAI champion (2020) and a mission-driven contributor for education and open innovations in ML, data science and AI.

Laiha Mat Kiah is a professor at Faculty of Computer Science and Information Technology, Universiti Malaya. Current research interests include Cyber Security, Secure Health Information Exchange and Blockchain Technology.

Katharina Miller, Master of Laws (Luxembourg), is a change agent using legal tools for ethics and integrity in innovation and technology, currently working for the H2020 project Path2Integrity. Katharina is an expert in corporate governance, corporate compliance, gender equality and human rights. She is a founding partner of 3C Compliance, member of the advisory board of PANDA GmbH and supervisory board member of ASI GmbH and Vinoselección. She used to be president of the board of directors of the investment fund KOKORO and member of the sustainability panel of Telefónica. Katharina is the President of the European Women Lawyers Association (EWLA) as of November 2017. She is founding member of the weekly newsletter "Leading with AI". Katharina is a sought-after speaker in the field of diversity and corporate governance and Incoming Head of the EU delegation of the G20 / W20. In 2017 and 2018 she was voted as one of the Spain's Top 100 Women Leaders. She is also an Adjunct Professor at the IE Law School and works as a European Commission Reviewer and Ethics Expert. She is an Advisory Board member of the Berkeley Center on Comparative Equality and Anti-Discrimination Law. Katharina is co-editor of the collective book "The Fourth Industrial Revolution and its Impact on Ethics – Solving the Challenges of the Agenda 2030" which has been published in January 2021 with the editorial Springer.

Cynthia Montaudon-Tomas is a full-time professor and a business consultant at the Business School at UPAEP. Since September 2018, she became head of the Observatory on Competitiveness and New Ways of Working, which deals with numerous issues related to social problems and creates awareness about current and future social needs. She has obtained a Post Doctorate Certificate In Organizational Leadership from Regent University in Virginia, USA; a Ph.D. in Strategic Planning and Technology Management from UPAEP, in Puebla, Mexico; a Ph.D. in Business from the University of Lincoln in Lincolnshire, UK, and three masters degrees: one in Quality Engineering from the Universidad Iberoamericana in Puebla, Mexico, another in Communication and media from the Jean Moulin, Lyon II University in Lyon, France, and the last one in Business Administration from the Tec de Monterrey in Mexico.

Ingrid Pinto López is a university professor, researcher and consultor at Business School of UPAEP University, recognized by the National System of Researchers (SNI Conacyt). Ph.D. in Strategic Planning and Technology Management. Currently coordinator of the competitiveness observatory and new forms of work, coordinator of the international arbitration of the Latin American Association of Accounting and Administration Schools ALAFEC, member of Barcelona Economics Network and the Illustrious Latin American Academy of doctors.

Samia Rizk is an Emirate Professor of Clinical Pathology at Cairo University School of Medicine in Egypt, with a special experience in hematology and hematopathology, medical education, and research ethics. She had 20 years of work experience in clinical laboratory practice, research, and management in the health field, Pediatric hematology, Consultant of laboratory hematology for FUO (fever of unknown origin) cases at Abbassia Fever Hospital-NAMRU-3. She is a member of the Cairo University IRB and has contributed to the planning and teaching of postgraduate courses in research ethics and participated in many national and international ethics meetings and workshops. Global research ethics is one of her special interests. She participated in 2 global research courses namely; the "Ethical Issues In Global Health Research (EIGHR)" on September 23-26, 2019 at Harvard Medical School Boston, MA, and the 2020 Global Bioethics Initiative course online Dr. Rizk was certified as an IRB professional (CIP) in, 2015 with recertification in 2018, and 2021. In addition, she is also a certified Research ethics trainer through "the Research ethics training" program at the University of Maryland, in May 2009. She authored a book's chapter: "Risk-benefit Assessment in Clinical Research Practice" in the: "Ethics in Research Practice and Innovation" book, published by IGI Global, 2020. (/book/ethics-research-practice-innovation/199086), 2019, pages 148-170. DOI: 10.4018/978-1-5225-6310-5.ch008.

Susan E. Zinner is a Professor in the School of Public & Environmental Affairs at Indiana University Northwest. She obtained her undergraduate degree in English literature cum laude from Bellarmine College (now Bellarmine University) in Louisville, Kentucky, her master's degree in journalism from Northwestern University in Evanston, Illinois and her law degree and master's degree in health administration both from Washington University in St. Louis, Missouri. She teaches undergraduate and graduate classes in health administration, law and ethics. She publishes and speaks at international conferences on bioethics, with an emphasis on pediatric and vulnerable populations. She served as Faculty President on her campus 2016-20.

Index

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

V

Y