Challenges of Digitalization: Algorithms and Fake News

IT-489-A: Project Report

Ina Budina

Marymount University

Dr. Ermicioi

07/23/2021

Table of Contents

List of Tables

List of Figures

Abstract

Digitalization and Artificial Intelligence (AI) are all facts today. However, the bias embedded in algorithms presents important challenges, such as the potential to produce fake information, which in turn can allow for foreign interference in the democratic political process and could be used to polarize society. For example, AI has facilitated the Russian interference in the United States presidential elections, while algorithms are in part responsible for the spread of fake news, social media has been used as a tool for technology warfare. This project offers potential policy solutions to help online users classify fake news, by enhancing their understanding of how algorithms work, and by enhancing the ability of social media to protect its users from disinformation campaigns.

Keywords: Digitalization, artificial intelligence, algorithms, risk management

Challenges of Digitalization: Algorithms and Fake News

Introduction

Many technological advancements have changed our lives in the last century and this trend certainly accelerated because of COVID-19, when many businesses and the whole world had to rapidly digitalize in a matter of weeks. FinTech revolution, digitalization, the employment of AI and incorporating voice recognition are all facts today. In this modern age of technology, digitalization holds the promise of endless capabilities for anything. With digitalization, comes AI, or 'Artificial Intelligence' which is used so often nowadays, as can be seen with different companies like Google, Siri or Samsung. Of course, for these processes to function, algorithms are needed and those can be described as a specific number of rules that should be followed in a problem-solving process done by a computer.

At the same time, algorithms present important risks and can be very harmful to society. For instance, they have the ability to produce disinformation or 'fake news', multiplying them through the system, filter and present only biased information to anyone's social media apps, with the potential to use algorithms as a warfare tool, and allow foreign interference in politics and in nearly any sphere of our lives today. Algorithms can therefore be harmful, e.g. lead to increased social polarization, while security breaches can harm the economy, consumers, and society.

In my project, the main research questions will center around the following: (i) What are the risks and challenges of algorithms and digitalization? (ii) How significant are these risks? and (iii) What are the policy options for a robust risk management strategy?

I plan to investigate these questions using several approaches. First, a critical literature review, along with several concrete case studies will be used to analyze key risks which if materialized can cause significant harm to society. I will also research and present information on potential losses from the materialization of key risks. Finally, the novelty of the capstone project will be to present various policy options for risk management strategies to minimize risks related to algorithms, including the role of government regulatory institutions, consumer protection bureaus, self-regulation by tech companies.

Project Proposal

OVERVIEW

Digitalization made it possible for many businesses to stay afloat, for employees to work from home, and for many students to continue to study from home. Moreover, the spread of online investment platforms like Robinhood, which did not charge any fees for investing in the stock market, democratized the financial industry, as even high school students with a little cash could invest in the stock market. Furthermore, telemedicine began flourishing, which was a cheaper and safer way to provide the service and reach more people.

However, algorithms and digitalization also pose significant risks and challenges. For instance, they can multiply and filter 'fake news', aiding disinformation campaigns of malicious foreign interference, polarizing today's society. Algorithms can exploit vulnerabilities and enable security breaches with severe economic consequences, e.g. by harming vital US infrastructure, for instance in the context of the latest gasoline crisis, whereby the hackers asked for $5 million or the alternative was to leave consumers without gasoline. Severe data breaches, for instance in Ireland, led to personal data of Irish patients stolen by hackers that later on appeared online. In addition, even our democratic institutions and electoral process could be called into question, for instance in the context of the role of Cambridge Analytica and its unethical use of Facebook data for the purpose to sway voting preferences of Facebook users in the context of the 2016 US presidential elections. Most recently, the COVID19 tracing apps has also raised important ethical issues around peoples' rights for privacy, and the use of AI.

OBJECTIVE

In my capstone project, I would like to analyze challenges of algorithms and digitalization, particularly when done in such a disruptive way, and then come up with a policy plan for addressing the various challenges it presents, e.g. risks related to potential foreign interference in the political process, cybersecurity risks, as well as various ethical issues centering around the use of algorithms and AI, that may be harmful to society, e.g. their potential to polarize society and to discriminate against groups of individuals, race, gender, etc., as algorithms might embed such biases, while lacking transparency. This project will explore policy options to address key risks that algorithms may present, to ensure that digitalization could deliver benefits to all and help address key risks.

FACULTY CONTENT ADVISOR WITH SUMMARY OF QUALIFICATIONS

The nature of my project is research and as such there will be no faculty content advisor. My only advisor for this project will be Dr. Natalia Ermicioi, who holds the position of an Information technology professor at Marymount University. Given her knowledge and extensive experience with conducting research projects, she would be a great advisor and would help stay focused, manage deadlines, accordingly, securing writing assignments on time, and realize project ideas for scaling up.

PROJECT PLAN

The plan for my project is sufficiently detailed, breaking down the project into key tasks and outputs, and ensuring that the timeline is within the deadlines as required by my advisor. Specifically, Table 1 presents the key tasks, broken down in the following elements: (i) the choice of capstone topic; (ii) the project proposal and a plan for the project; (iii) the literature review, along with gathering information - scholarly articles from JSTOR, book chapters, government reports, as well as companies/ reports, also summarizing each source; (iv) structure of the report, as well as a breakdown for the production of each chapter; (v) production of the draft report; (vi) the main report presentation, and (vii) the final report.

Furthermore, to make sure that I make satisfactory progress on the report, I have also created a timeline for each separate task related to the production of my main report, using schedule/GANTT chart (Figure 1). Each row in the chart presents various output under my project, while each column shows the overall timeline for the project in weeks. The blue bars under each task show the duration of each task, and when is the due date. Since we have our deadlines each Friday, each colored cell represents a work week.

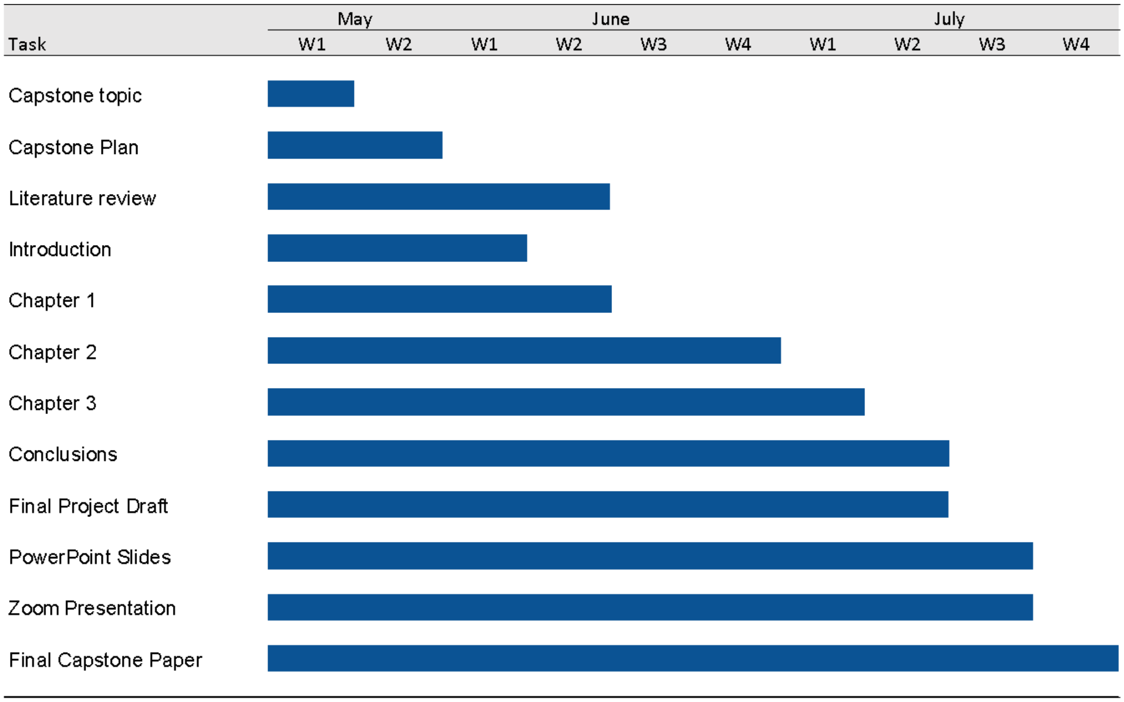
I have already selected my project topic, prepared a proposal with a project plan and timeline, researched the literature, produced a summary for each source, and prepared a literature overview. I have also drafted all the individual chapters and put them together in a draft project document.

Table 1 Project Plan

Start Date	Expected End Result	Task	Description	Level of Effort
05/15/2021	Capstone topic Done	AI and conflicting information	Exploring the effects of AI and information rabbit holes	9
05/28/2021	Capstone Plan Done	How fake info is created	Looking into how and why fake information is created and affects people	9
06/04/2021	Critical Literature Review	Summarize key issues	Review 10 professional papers	9
06/04/2021	Introduction	Complete the chapter	Why addressing issues of AI and fake information and creating a solution	9
06/18/2021	Chapter 2 Literature Review	Draft the chapter	Review 10 professional papers	9
06/25/2021	Chapter 1: Algorithms	Complete the chapter	Algorithms: definitions, examples, how do they operate.	
06/25/2021	Chapter 2: Challenges of Algorithms	Complete the chapter	Potential information issues: fake news, data privacy, national and international impacts	
07/02/2021	Chapter 3: Policy Solutions for Detecting and Mitigating Risks	Complete Novelty chapter	Possible solutions: 1.give control to users to detect fake	

			news, 2. role of social media companies	
07/09/2021	Conclusions	To complete	Addressing issues of AI and fake information and creating a solution	
07/09/2021	Final project draft	To complete and submit	Nearly final shape with all chapters included	
07/16/2021	PowerPoint Slides	Live zoom presentation		
07/23/2021	Final Project	To submit		

Figure 1Project Planning (Gant Chart)



RESOURCES

In developing, organizing, and completing the project different resources will be used. Main source of information that would be used are scholarly articles. To secure credible information, the database JSTOR, not-for-profit service, will be used. This trusted digital archive has a wide choice and variety of respected academic articles that can help scholars, researchers, and students to build on selected ideas that are beneficial to the project. At least ten articles will be used.

In addition other sources like Google Scholar and Google online search will be used to research topics and gain useful information, particularly with regards to checking the webpages of some agencies that provide guidelines with regards to addressing some risks related to algorithms, e.g. consumer protection bureaus, etc.

A third source that can be beneficial to deepen further the required knowledge will be documentary movies related to the selected topic. The documentary titled “The Great Hack”, 2019 and “Coded Bias” gives some very useful information and a very detailed insight on algorithms, e.g. explaining how disinformation commonly can occur, which can be used in developing the project, or highlighting the implicit biases embedded in algorithms.

The fourth source will include some articles from different technology related magazines and IT company publications. Given that the credibility of these sources is determined, useful facts and ideas will be used while working on the project.

EXPECTED OUTCOMES

The expectation is that this research project will lead to an interesting and informative paper, which will give insights about the use of AI considering benefits and risk of application of the information technology. It will summarize past experience, provide valuable information about key risks related to the use of algorithms and AI, distill pros and cons of different options for reforms, and offer a framework for managing algorithmic and AI risks that can be used by businesses, such as social media companies, but also the public sector to at least minimize and avoid risks, such as fake news.

KNOWLEDGE BEING APPLIED

As a senior in Information technology major at Marymount university I have a challenging academic program, where I have taken many classes related to Networking, Cybersecurity, Computer Technology and others. The knowledge I have gathered during the past three years is based on the foundational courses like a number of business and economic courses, cyber ethics, philosophy, and religion, as well as more technical courses like Python, Web Development, Database management. In this respect I will use many aspects of the technology knowledge I gained during my education along with project and time management skills I have developed previously. Most useful skills will be ability to scale and focus on the selected topic and bring my own opinion and vision, to always think of risks and vulnerabilities, and design risk management strategies. Also, the ability to set strict deadlines and goals is crucial.

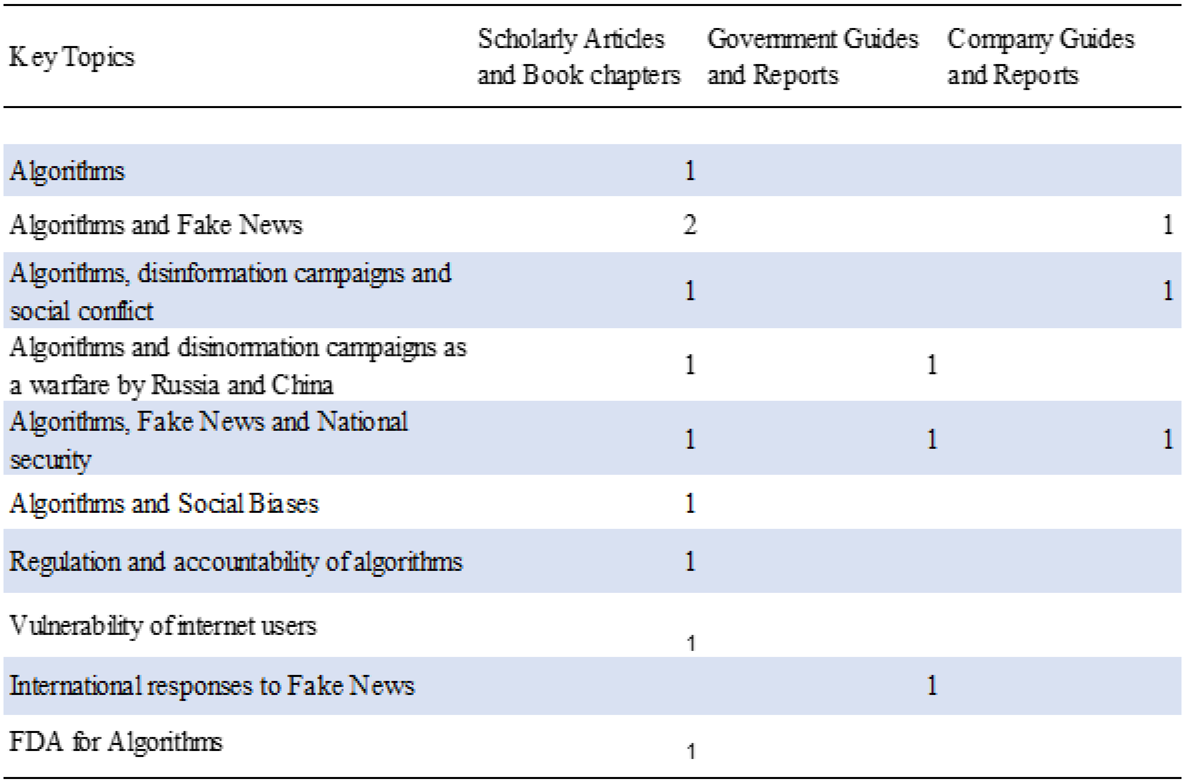
Chapter 2: Literature Review

INTRODUCTION

The goal of the literature search was to find scholarly articles with a rich discussion on algorithms, definition, and risks and challenges embedded in them. For instance, there are several articles that focus on algorithms and fake news. Other topics include the ability of algorithms to propagate disinformation campaigns and induce social polarization and conflict. There were also several publications - articles and reports on foreign interference via disinformation campaigns and risks and challenges posed for national security. Many articles discuss the algorithm biases and how algorithms can discriminate as well, simply due to the way they are written, reflecting the biases of people who created them. Finally, a few reports and papers also focused on potential solutions and policy options for risk management, including how to address vulnerabilities of internet users, how to make algorithms more accountable, and more transparent about biases embedded in them. There were also a few sources about proposals on how to counter fake news, about the need of tech companies to self-regulate, potential changes in legislation, about the need of "FDA for Algorithms", and about ways to enhance the literacy of the population about the algorithms.

The first section summarizes the literature, including the description of the researched topic. The second section discusses the strategy search, including the strategy of research of papers and collecting information along with keywords that have been used. Next, each source is summarized and briefly analyzed as a related referenced source for the research project. The last section includes conclusions of the literature review and a short reflection on the availability and credibility of the researched information and usefulness of the academic articles for the project research.

Table 2. Key topics and resources.



Key Topics	Scholarly Articles and Book chapters	Government Guides and Reports	Company Guides and Reports
Algorithms	1		
Algorithms and Fake News	2		1
Algorithms, disinformation campaigns and social conflict	1		1
Algorithms and disinormation campaigns as a warfare by Russia and China	1	1	
Algorithms, Fake News and National security	1	1	1
Algorithms and Social Biases	1		
Regulation and accountability of algorithms	1		
Vulnerability of internet users	1		
International responses to Fake News		1	
FDA for Algorithms	1		

SUMMARY OF TOPIC

Given the accelerated development and use of algorithms and artificial intelligence in every sphere of our business and personal life, the literature search focused on potential challenges and risks from the use of algorithms and AI. According to “Unpacking Fake News” by Norman Vasu, who wrote about the potential of algorithms to spread ‘fake news’ can be used for campaigns aiming to destabilize the states by attempting to demolish modern society from the very inside by distorting real time information deliberately. According to the author, the algorithms are a powerful tool which can be used to deliberately sway the outcome of political elections or weigh the opinion polls on political matters, e.g. Brexit vote in the United Kingdom, which can undermine free democratic elections in this tech advanced era. Moreover, a number of articles also conveyed that using algorithms does not mean that we are dealing with the objective truth, that the mathematics behind the algorithms does not necessarily guarantee the truth, but rather, they reflect the biases of humans that have created algorithms, so they may still be ethically immoral. Algorithms, usually used by information intermediaries, or platforms, can embed a bias by promoting selective information to each user based on the way information is sorted, filtered, and ranked. Moreover, social networks that collect a vast amount of personal information, can have perverse incentive to trade this information, which in turn could be used in a ways that undermines democratic institutions and political process, allows for foreign interference in the election process, with the potential to sustain social conflict and polarize the society.

The second goal of my literature search was to find literature with proposals on how to address risks and challenges embedded in using algorithms and AI, which could help me think about the proper risk management strategy. For example, would more government intervention help address risk, should we nationalize internet companies? Most of the articles I have reviewed confirmed that this is not the solution; the way to go is to enhance transparency and accountability of the social networks for the algorithms they use, as well as incentives for self-regulation by the social network. There were also a couple of articles that argue that new regulation of AI should only be created for new social risks that cannot be dealt with existing legal and institutional frameworks. Finally, there were several papers that argued for FDA for algorithms and AI, and that big tech should create a Board of Directors with independent experts to ensure self-regulation, social responsibility, and accountability. Finally, from an economic perspective, a risk management strategy would require several elements: (i) collecting information and tools for adequately assessing risks from AI, (ii) public discussion of risks and designing a risk management strategy; (iii) elements of risk mitigation (e.g. regulatory, institutions, including the role of audit institutions for assessing risks related with AI), and (iv) preparing a strategy for absorption of the residual risk.

DESCRIPTION OF SEARCH STRATEGY

The literature search strategy is based on searching for articles and books from JSTOR, Google Scholar, ResearchGate, and other journal articles, as well as reports from government regulatory agencies. In addition, the topic selection and discussions have also benefited from several documentaries and a broader discussion by the public, particularly on the issue of Russian interference in the United States Elections, and most recent news of high profile hacking by criminal groups, demanding ransom, which have been a lot in the public domain lately. The search of my relevant sources often included words like “algorithms”, “fake news”, “Cambridge analytica”, “AI regulation”, “risks from using AI” and “AI and machine learning”.

SOURCE REVIEW

Allcott, H. and M. Gentzkow (2017). Social Media and Fake News in the 2016 Election. *The Journal of Economic Perspectives*, vol. 31, no. 2, 2017, pp. 211–235. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/44235006.zz

The increased importance of social media in producing fake news, particularly in the run-up to the 2016 elections and the implications for the voting patterns in 2016 U.S. elections is one example of challenges related to the use of algorithms (Allcott, 2017). This 2017 scholarly article from the Journal of Economic Perspectives is relevant for my research as it discusses issues surrounding social media, particularly as the media uses algorithms and targets individuals with fake news, skewed towards one of the political candidates ahead of the presidential elections. The algorithms filter the information and fake news, creating “filter bubbles" for users, without exposing them to different and contrarian perspectives. The authors also provide an estimate of the extent to which social media users have been exposed to fake news and discuss possible impact of fake news on the voting patterns during the 2016 US elections.

This article is relevant for my research as it highlights the potential negative effects of algorithms on society, particularly the ability to carry out democratic elections.

Keywords: Social media, fake news, elections

Asmolov, G. (2018). The Disconnected power of disinformation campaigns. *Journal of International Affairs*, vol. 71, no. 1.5, 2018, pp. 69–76. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26508120

Fake news campaigns are malicious as they can sustain social conflict and divisiveness in society (Asmolov, 2018). In this special issue of the 2018 Journal of International Affairs, Gregory Asmolov highlights the role of technology in generating fake news and weaponizing disinformation to create and sustain social divide and conflict. The article also offers ideas about policies to address such disinformation campaigns particularly in a conflict state.

Keywords: fake news, social discourse, conflict

Balkin, J. (2020). How to Regulate (and Not Regulate) Social Media. Essays and Scholarship. Knight First Amendment Institute at Columbia University. Retrieved from How to Regulate (and Not Regulate) Social Media | Knight First Amendment Institute (knightcolumbia.org).

More government regulation will not adequately address challenges with algorithms and social media; it would make things worse. Balkin argues for smart regulation, aiming to create “incentives for social media companies to be responsible and trustworthy institutions” (Balkin, 2020). This article is directly relevant for the policy section of my research project, how to design proper risk management strategies.

Keywords: regulation; social media; incentives

Callahan, G. (2021). What are the potential risks of AI? Jan 5, 2021. Blog. Artificial Intelligence. https://www.rev.com/blog/what-are-the-potential-risks-of-artificial-intelligence.

The accelerating development and use of AI have enormous potential, but many tech leaders and experts (Stephen Hawking, Bill Gates or Elon Musk) are warning about their potential risks (Callahan, 2021). The author warns that AI could lead to major bias based on race or gender, inequality, human job loss and physical harm and examines some of the most common concerns about AI and the risks it poses.

Keywords: artificial intelligence, potential AI risks

Hunt, R., and F. McKelvey (2019). Algorithmic Regulation in Media and Cultural Policy: A Framework to Evaluate Barriers to Accountability. *Journal of Information Policy*, vol. 9, 2019, pp. 307–335. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/10.5325/jinfopoli.9.2019.0307

Algorithms are defined as an automated decision-making process, though modern-day algorithms are also self-taught because of modern day machine learning (Hunt, 2019). In this 2019 scholarly article, Robert Hunt and Fenwick McKelvey from Concordia University explain how internet platforms - or the information intermediaries - are reliable on algorithms that will sort, filter, rank and promote selective information to individual users. The authors discuss controversies related to algorithms, e.g. the embedded biases associated with developing and deploying such algorithms in different spheres, which are used to classify, sort and promote information content. The article is directly relevant for my research project, and particularly the discussion about the need of algorithmic accountability as a new area of regulation and about calling for measures to enhance algorithmic accountability.

Keywords: algorithms, accountability, cultural policy, discoverability, artificial intelligence

Krishna, D., N. Albinson, and Y. Chu (2017). Managing Algorithmic risks. Safeguarding the use of complex algorithms and machine learning. Deloitte report. Retrieved 3 June 2021, from Managing algorithmic risks (deloitte.com).

Kirby, R., K. Vlaskova, J. Kolenick, and P. Kubala (2018). Online habits of the fake news audience: the vulnerabilities of internet users to manipulations by malevolent participants. *Geopolitics, History, and International Relations*, vol. 10, no. 2, 2018, pp. 44–50. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26802338

A summary of literature highlights internet users’ vulnerability to manipulations and using fake news as a warfare tool (Kirby, 2018). Based on data from various sources, this survey paper by Raffaella Kirby and others estimates the most likely sources of fake news in the US and the share of adults that are confident that they can spot fake news. The article is relevant, as the discussion is based on a literature review from many other studies, and it is objective as it uses an empirical-based approach.

Keywords: online habit; fake news audience; Internet user; manipulation; vulnerability

Levy, R. (2017). Taking Aim at Biased Algorithms. *Math Horizons*, vol. 25, no. 1, 2017, pp. 5–7. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/10.4169/mathhorizons.25.1.5

Using algorithms, or automated processes does not guarantee truthfulness and trustworthiness of information (Levy, 2017). This scholarly article in Math Horizons, by Rachel Levy, argues that mathematics behind the algorithms doesn't necessarily guarantee the truth and that behind the mathematics and the code of algorithms there are humans, each one of them with their own biases, reflected in the codes. Historical biases also reflect a repeat of past mistakes, so mathematics is often used to cover up and defend things that are wrong and even immoral.

Keywords: algorithms; bias

Metaxas, P. T., and E. Mustafaraj (2012). Social Media and the Elections. *Science*, vol. 338, no. 6106, 2012, pp. 472–473, retrieved from www.jstor.org/stable/41703780

Social media is the main source of manipulation and it has been used many times to affect political decisions during elections (Metaxas, 2012). The authors argue that “google bombs”, “twitter bombs”, or “prefab tweet factories” and online search ads are just a few ways to manipulate elections by influencing voters with purposely selected information and that social media is a very low cost and effective tool for manipulating and spreading select and often fake news. Social media is often used to predict political elections and results from different events, given that it offers widespread and easy access to its users, though these predictions are controversial.

Keywords: social media; elections

Prier, J. (2017). Commanding the Trend: Social Media as Information Warfare. *Strategic Studies Quarterly*, vol. 11, no. 4, 2017, pp. 50–85. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26271634.

This scholarly article examines two case studies - one related to the Islamic State, and the second one - to Russia election interference to provide evidence for how social media has been successfully used for propaganda (Prier, 2017). This article is relevant for my research, as it shows the potential for using algorithms and social networks to spread propaganda as a warfare and the potential for many more countries to build such propaganda campaigns on social media.

Keywords: propaganda; social media; information sharing

Reed C. (2018). How should we regulate artificial intelligence? Philosophical Transactions. Royal Society Publishing. A 376: 20170360. Retrieved from http://dx.doi.org/10.1098/rsta.2017.0360

The implementation of one general system for AI regulation is not appropriate and it is too early to attempt; instead, regulation should work incrementally, based on already existing legal and regulatory schemes (C., 2018). While in many cases the current legal system might be efficient and can solve problems, AI creates new societal risks and if AI expansion creates specific risks which current law and regulation cannot cover, then new regulation would be required. Good regulation should therefore remain focused on those new unknown risks brought by AI solutions.

Keywords: artificial intelligence, machine learning, law, regulation, transparency

Rieder, B., and J. Hofmann, (2020). Towards platform observability, Internet Policy Review, Journal on internet regulation, 18 Dec 2020 DOI: 10.14763/2020.4.1535, Retrieved from Towards platform observability | Internet Policy Review.

This is a research article, published in a peer reviewed open access journal. It is relevant for my research project, as it proposes some broad principles as regulatory guidelines for making platforms more accountable to their users, which is an important component of designing a risk-management strategy for algorithms.

Keywords: platforms; accountability; regulation

Roese, V. (2018). You Won't Believe How Co-Dependent They Are: Or: Media Hype and the Interaction of News Media, Social Media, and the User. *From Media Hype to Twitter Storm*, edited

by P. Vasterman, Amsterdam University Press, Amsterdam, 2018, pp. 313–332. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/j.ctt21215m0.19

Social media, regular news media, and regular users are tightly connected and the change of each one, the social or regular media, or their uses can lead to changes in the other two categories (Roese, 2018). Vivian Roese argues that social media has become the most dynamic element in those relationships, as it is the main driver of news and communication in recent years, hence influencing profoundly both regular news media and the public. The author is discussing the interconnection, benefits and risks for those three elements, as there is also a darker side of social media - the cost of how destructive it can be when fake news and lies are taking over in the social media space.

Keywords: interaction of news media, social media, and the user

Vasu, N., B. Ang, T. Teo, S. Jayakumar, M. Faizal and J. Ahuja (2018). Fake news: national security in the post-truth era. *S. Rajaratnam School of International Studies*, 2018, pp. 18–25, Retrieved 29 May 2021, from www.jstor.org/stable/resrep17648.8

Countries have used different approaches that are dealing with countering fake news (Vasu N. B.-A.-t.–2., 2018). This report surveys different approaches based on the nature of fake news and specific country considerations. The report describes individual countries' approaches and multilateral efforts like for example the EU where they have used volunteers for fact checking. Since most of the disinformation usually originates on social media platforms, the report argues for self-regulation by technological companies, giving as an example Facebook, which is using fact check to deal with fake news. The report also includes policy implications for combating fake news.

Keywords: fake news; national security

Vasu, N., et al. (2018). Unpacking Fake News. ***S. Rajaratnam School of International Studies*****, 2018, pp. 5–9, Fake news: national security in the post-truth era. Retrieved 29 May 2021, from** www.jstor.org/stable/resrep17648.5

The use of ‘fake news’ aims at distorting real time information to destabilize the states by attempting to demolish modern society from the very inside (Vasu N. B.-A., 2018). According to “Unpacking Fake News” by Norman Vasu, a straightforward way to start up is by pivoting on elections or other political matters or events, which would create the largest influence, in this tech advanced era.

Keywords: fake news; propaganda; elections

Weissmann, S. (2019). How Not to Regulate Social Media. The New Atlantis No.58 (Spring 2019), pp. 58-64. Retrieved from How Not to Regulate Social Media on JSTOR.

There is evidence that more government regulation will not help challenges and risks posed by social media, and bots (Weissmann, 2019). This scholarly work discusses ways to help address challenges presented by social media and algorithms, specifically referring to the Facebook- Cambridge Analytica scandal and the interference in the United States election. This piece is relevant to my policy section, specifically, to design a risk management strategy.

Keywords: artificial intelligence, machine learning, law, regulation, transparency

REFLECTION ON THE AVAILABILITY OF INFORMATION

There is a lot of information about algorithms and use of artificial intelligence in many aspects of our life. There were plenty of available articles about the potential danger of using AI in collecting data and using it to manipulate information or creating fake news, when dealing with elections, other political decisions or polarizing the society. There was enough information analyzing data bias used in AI as a main factor of manipulation and creating misleading conclusions.

There was also a good amount of information about legal and ethical implications of using AI. This search supplies a very important aspect of extended use of algorithms and helps deeply analyze the moral of such technological advancement in our society. It further complicates the topic as an interrelated (moral technology) development dilemma of our society.

The most difficult part was to find information about elements of a good risk management strategy of Artificial Intelligence, given that articulating of such a strategy is still in an early state. At the same time, this is what has made this topic interesting as the challenge to our society and institutions is enormous and even our democracy is at stake. I plan to include a few more sources (scholarly articles) and look for more information, especially risk management from specialists and experts to further develop my research.

CONCLUSION

This chapter provided a thorough literature review, going over a number of sources, summarizing key issues learned from the literature review, and with a detailed description of the search process, including keywords used, as well as the search platforms. The review also included the gist of the discussion included in each of the sources that will be used as a basis for this project, and also explaining why I picked each one of these sources, how they help and contribute to the research included in this project. Finally, the review also highlighted areas where the literature was scant, with regards to my policy sections, which is the novelty part of the project.

Chapter 3: Project Novelty

ALGORITHMS

Algorithms drive almost any sphere of our lives today. They penetrate virtually all sectors of our economies, and they occupy a progressively large part of our free time, social connections, personal communications, entertaining, travelling, and even control heating and AC systems in our homes. For example, they help us get to our favorite shows and movies on Netflix, watch YouTube videos on a particular topic of interest, help us find the most convenient flight tickets or favorite vacation getaways, or buy goods and services from the convenience of our houses. However, with every click, we create a datapoint, allowing for the machine to gain an insight about our habits, tastes and psychological profiles with a precision that has never been possible in the past, allowing various businesses to target their consumers with incredible accuracy.

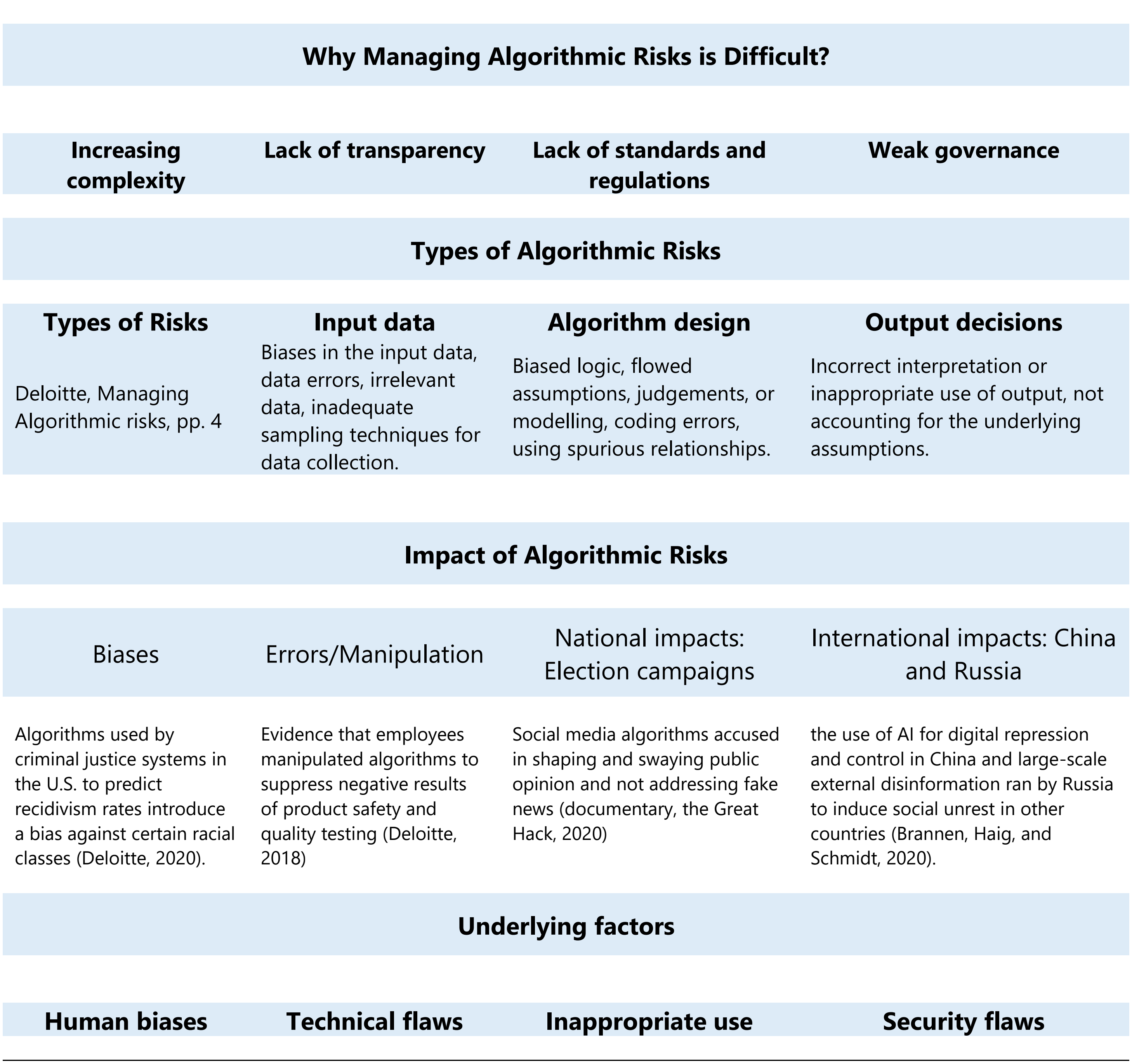
But what are the Algorithms? Algorithms are a set of rules or processes that define step by step instructions to solve a problem. They define how the work should be executed to achieve expected results and outcomes, how to solve specific problems, and what operations to execute. Early algorithms were just a set of basic commands programmed to perform certain tasks. However, with Artificial intelligence (AI) and machine learning, algorithms are becoming increasingly complex, with capability to “self-learn” based on training datasets. Advancements in deep learning techniques enables algorithms to solve complex problems and make predictions, based on sophisticated modelling techniques, and powered by big data.

CHALLENGES OF ALGORITHMS: RISKS AND VULNERABILITIES

What are the risks and challenges created by algorithms? As highlighted in the Deloitte report by (Krishna, 2017), advances in AI and machine learning technology, along with powerful big data created by their users present important challenges and risks. Some of the challenges, highlighted in the literature include increasing complexity, lack of transparency around algorithm design, lack of standards and regulations, and weak governance (Table 2). These challenges pose new risks to organizations, public institutions, and society, such as the prevalence of fake news, biases, errors, lack of privacy, and malicious acts (Callahan, 2021). For instance, Kirby analyses the vulnerability of internet users in the United States to manipulations and their ability to spot fake news (Kirby, 2018). The cost of how destructive fake news can be are increasing in recent years, given that social media has become the main driver of news and communication in recent years, influencing profoundly both regular news media and the public (Roese, 2018). Moreover, disinformation campaigns have also been found to have a destructive power with the potential to break the social fabric of our modern society (Asmolov, 2018).

Algorithmic designs are also prone to biased logic, inaccurate beliefs or opinions, unfitting modeling techniques, mistakes in coding, and sometimes they even may identify spurious patterns in training data. The materialization of such risks could have significant national and international impacts. For instance, literature highlighted the risk of using algorithms and platforms as warfare, highlighting two examples (one related to the Islamic State, and another one - to Russia’s interference in the 2016 US elections) as evidence about using social media as propaganda (Prier, 2017) (Metaxas, 2012). In China, there is also a proliferation of military research establishments, such as Beijing Institute of Technology which aims to generate an army of high-tech soldiers, equipped to operate AI weapon systems ranging from tanks, drones and submarines to microscopic high-tech robots.

Table 3 Algorithmic Risks: Causes and Impacts



Why Managing Algorithmic Risks is Difficult?			
Increasing complexity	**Lack of transparency**	**Lack of standards and regulations**	**Weak governance**
Types of Algorithmic Risks			
Types of Risks	**Input data**	**Algorithm design**	**Output decisions**
Deloitte, Managing Algorithmic risks, pp. 4	Biases in the input data, data errors, irrelevant data, inadequate sampling techniques for data collection.	Biased logic, flowed assumptions, judgements, or modelling, coding errors, using spurious relationships.	Incorrect interpretation or inappropriate use of output, not accounting for the underlying assumptions.
Impact of Algorithmic Risks			
Biases	Errors/Manipulation	National impacts: Election campaigns	International impacts: China and Russia
Algorithms used by criminal justice systems in the U.S. to predict recidivism rates introduce a bias against certain racial classes (Deloitte, 2020).	Evidence that employees manipulated algorithms to suppress negative results of product safety and quality testing (Deloitte, 2018)	Social media algorithms accused in shaping and swaying public opinion and not addressing fake news (documentary, the Great Hack, 2020)	the use of AI for digital repression and control in China and large-scale external disinformation ran by Russia to induce social unrest in other countries (Brannen, Haig, and Schmidt, 2020).
Underlying factors			
Human biases	**Technical flaws**	**Inappropriate use**	**Security flaws**

Source: Deloitte, 2018, Managing algorithmic risks. Safeguarding the use of complex algorithms and machine learning.

These examples also show the potential for many more countries to build such propaganda campaigns on social media. A few articles highlight the significant national impact of social media algorithms in terms of swaying public opinion and the 2016 US elections, creating a potential to compromise democratic institutions by “creating opinion echo chambers and failing to clamp down on fake news” ((Krishna, 2017), (Vasu, International Responses to Fake News., 2018), (Allcott, 2017)).

A lack of a good risk management system for algorithmic risks may undermine all positive effects of the use of algorithms. Because of their use as a strategic decision making, algorithms may impose financial, regulatory, and strategic risks for many organizations if any errors or curtain vulnerabilities in algorithms appear. Efficient management of algorithmic risks requires an innovative approach, combining strong risk management at the individual organization level, with regulatory reforms and adapting new leading practices. A successful algorithmic risk management at the individual organization level, requires setting risk management frameworks to evaluate potential issues with algorithmic deployment, evaluate the potential impact of algorithms risks, enhance senior management’s understanding of algorithmic risks, and set a governance structure to identify, assess, and manage potential algorithm risks. At the same time, the organizations are not alone in their efforts for managing algorithmic risks. Many researchers, regulators, lawmakers, and advocacy groups advocate for the need of reforms to improve risk management standards to match the increasing complexity in algorithmic use. For instance, efforts of fact checking by volunteer groups have became an important tool in dealing with fake news in a number of countries, including in EU (Vasu, Fake news: national security in the post-truth era, 2018)

MANAGING ALGORITHMIC RISKS: OPTIONS FOR REFORMS

This section focuses on options for reforms, including on pros and cons of various proposals on how to address risks and challenges embedded in using algorithms and AI, in an effort to identify elements of a proper risk management strategy.

Mitigating AI risks: Role of Regulation

A key question to policy makers today is how to adjust government regulation in the face of growing and often not-well understood algorithmic risks. For example, would more government intervention help address these risks? Should we nationalize all the social platforms that collect and distribute information and news? Most of the articles I have reviewed confirmed that this is not the solution, and there are several reasons why more government intervention would not solve the problem. More intrusive government interventions and larger print of the state is likely to pose new threats, including the risk of becoming a surveillance state, the risk of stifling innovation and private investment in technologies, as well as increasing inefficiencies and corruption.

Staying with government regulation, another issue debated in the literature is whether a new omni code for regulating AI in every sphere will help address these risks (Rieder, 2020)). Some articles argue in favor of creating an agency equivalent to the Food and Drug Administration (FDA) for approval of algorithmic adequacy, before they can be used by businesses (Tutt, 2016). Others argue for the need to develop an omni-regulation to regulate all the algorithms employed across all the industries and businesses (Rieder, 2020). However, most articles argue against such an approach (Weismann, 2019), citing the potential for conflicting regulations, e.g. banking and financial regulation may contradict the AI regulations that might also apply to banking and financial sectors, with the potential to create regulatory uncertainty for businesses operating in the relevant spheres. An emerging solution, proposed in the literature, is to create new AI regulation only for the new social risks that cannot be dealt with by the existing legal and institutional frameworks (Reed, 2018). For instance, the need for algorithmic accountability appears in a new area of regulation, including measures and standards for enhancing algorithmic accountability (Hunt, 2019).

Mitigating AI risks: Role of Governance

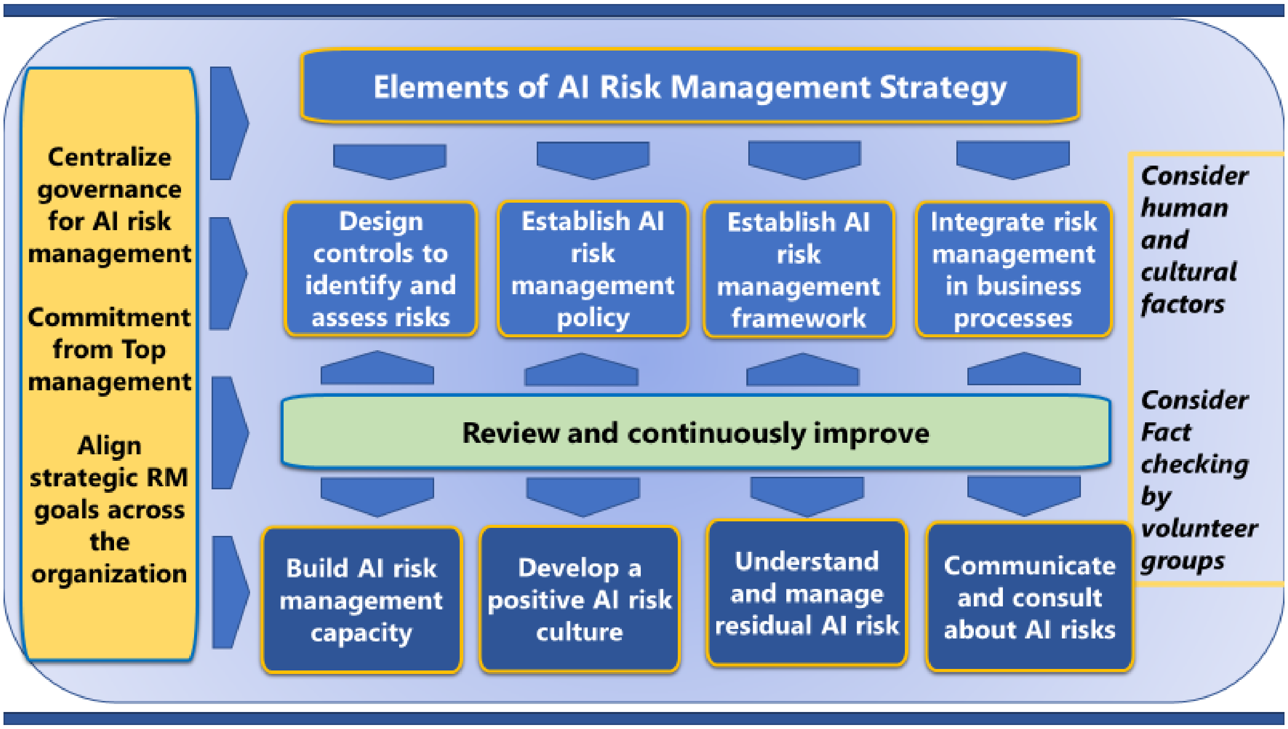
What are the options for reforms to minimize algorithmic risks? Literature argues about the need to enhance transparency and accountability of the social networks for the algorithms they use and improve incentives for self-regulation and committing to a morally ethical use of algorithms, including by social networks (GAO, 2018). For instance, Rieder and Hofman, 2020, propose a set of broad principles as regulatory guidelines for making platforms more accountable to their users, which is a key component of designing a risk-management strategy for algorithms.

Several papers argue that companies should establish a centralized body to manage algorithmic risk, adopt principles about spheres where algorithms can and cannot be applied, and set risk management policies, including for enhancing governance, accountability and checks and balances for companies that use algorithms. Some argue for the need to implement a smart regulation to enhance companies' incentives to self-regulate (Balkin, 2020). For instance, a requirement to include independent experts in the Board of Directors for companies using such algorithms can help enhance self-regulation, social responsibility, and accountability by overseeing companies' policies with respect to algorithms.

Elements of Risk Management Strategy for Algorithmic Risks

In response to raising algorithmic risks, organizations would need to adopt increasingly more sophisticated risk management strategies, build digital capacity, employ more technology-intensive tools and analytics to better assess AI risks, make their internal control system fit to detect AI risks, and strive to adhere to high governance, transparency, and moral ethics standards with respect to algorithms they use (Albinson et al., 2019). A stylized presentation of the model for algorithmic risk management is presented in Figure 2. The model brings together various elements of an algorithmic management strategy, including processes, analysis, institutional elements and for human and cultural factors.

Figure 2. Elements of Algorithmic Risk Management Strategy.


Elements of AI Risk Management Strategy
Centralize governance for AI risk management
Commitment from Top management
Align strategic RM goals across the organization
Design controls to identify and assess risks
Establish AI risk management policy
Establish AI risk management framework
Integrate risk management in business processes
Review and continuously improve
Build AI risk management capacity
Develop a positive AI risk culture
Understand and manage residual AI risk
Communicate and consult about AI risks
Consider human and cultural factors
Consider Fact checking by volunteer groups

Managing AI risks should be grounded in traditional risk management practices, while recognizing the complex nature of the algorithms, particularly the AI-based ones. From an economic perspective, a risk management strategy would require several elements: (i) Design proper governance arrangements to centralize AI risk management functions in companies, (ii) Design proper controls mechanisms, based on adequate information, and employing algorithms and analytics to identify and assess risks from AI, (ii) Promote collaboration and exchange of information on AI risks within the organization; (iii) Develop a risk management policy, including on elements of risk mitigation (e.g.

regulatory, institutions, including the role of audit institutions for assessing risks related with AI), (iv) Develop a framework for risk management that incorporates the regulatory and policy environment and the specifics of AI risks, and (v) Prepare to manage residual AI risk. Based on this stylized model for algorithmic risk management, organizations can use the following check list to assess their readiness to manage algorithmic risks (Box 1). A key part of their efforts also needs to focus on capacity development, and constantly improving the AI risk management systems.

Box 1: Policy Elements of Risk Management Strategy for AI

- Element 1: Centralize governance for managing AI risk
- Element 2: Design controls to identify and assess AI risks
- Element 3: Establish an AI risk management policy
- Element 4: Establish an AI risk management framework
- Element 5: Embed systematic risk management into business processes
- Element 6: Develop a positive AI risk culture
- Element 7: Communicate and consult about AI risks
- Element 8: Understand and manage residual AI risks
- Element 9: Build and Maintain AI risk management capacity
- Element 10: Review and continuously improve AI risk management

Tech companies, particularly those focusing on social media should have the technologies, and systems in place to detect inaccurate information and to minimize risks before they escalate (Albinson, 2019). In this respect, continuing education and enhancing the digital ability of the workforce should be a key priority for both the public and private sectors. Developing such ability is especially important for key control functions, such as audit teams, that would need to use high-tech tools to test or audit algorithms.

Chapter 4: Conclusions

This project researched and analyzed challenges and risks related to the wide-spread use of Artificial Intelligence and more specifically employment of the algorithms in various industries and organizations and highlighted the elements of a risk management framework that can be adopted to manage and address the algorithmic risks. The technological advancement, increasing capacities in programming and further development of new innovations led to a rapid digitalization of many industries and activities. The massive scale of internet use and the key role of social media in public life created an environment for exponential deployment of AI and algorithms in communications, news creation and broadcasting, while collecting vast information data from many users. Financial industry has changed radically with digitalization. Education system offers unlimited opportunities for online learning and many companies introduced distance work and learning. Along with many advantages, the use of AI and algorithms also pose new challenges and risks. Creating and spreading fake news, manipulation, misinformation, and election interference are examples of negative effects caused by often inappropriate use of AI and algorithms. Because of their use as a strategic decision making, algorithms may impose financial, regulatory, and strategic risks for many organizations if any errors or curtain vulnerabilities in algorithms appear. Therefore, analyzing literature is important for identifying elements of policy reforms and ideas about risk management and optimizing the regulation.

As the literature review suggests, the use of algorithms poses several new risks, including risks from embedded biases such as programmer errors, previous mistakes, inadequate models, or intentional efforts to produce wrong results and misinterpretations. Real examples of such risks are intentional spreading of selective information to misinform internet users, developing fake news and using previously collected user information to form and change opinions of users. Additionally, understanding the use of bots and their role to spread intentional information is critical to educate users how to avoid unethical behaviors and actively promote moral standards.

In view of many complexities embedded in algorithms and AI, efficient management of algorithmic risks requires innovative approaches, combining strong risk management frameworks at the individual organization level with regulatory reforms and adapting new leading practices. A successful algorithmic risk management at the individual organization level, requires setting risk management frameworks to evaluate potential issues with algorithmic deployment, evaluate the potential impact of algorithms risks, enhance senior management’s understanding of algorithmic risks, and set a governance structure to identify, assess, and manage potential algorithm risks.

The novelties of this project are that: (i) it took stance on the current debate on options for regulatory reforms to manage AI risks, and (ii) it proposed key elements of risk management strategies that could be adopted by the individual organizations to better manage AI risks. This research project has summarized the debate on options for regulatory reforms to manage AI and algorithmic risks and advocated for regulatory reforms to enhance transparency and governance of AI and algorithms. Furthermore, the project also put forth key steps that private and public organizations would need to take to implement or enhance traditional risk management frameworks to ensure a proper management of AI risks.

In conclusion, the development of ethical and moral standards of AI usage and proper management of the new AI and algorithmic risks require joint push form all participants, companies, regulatory authorities, civic and voluntary organizations, and advocacy groups. Only joint efforts will produce better results and help developing a successful risk management of the new AI risks.

References

Albinson, N., C. Thomas, M. Rohig, and Y. Chu (2019). Future of risk in the digital era. Transformative Change. Disruptive Risk. Deloitte report. Retrieved 3 June 2021, from Future of risk in the digital era (deloitte.com).

Allcott, Hunt, and Matthew Gentzkow (2017). Social Media and Fake News in the 2016 Election. *The Journal of Economic Perspectives*, vol. 31, no. 2, 2017, pp. 211–235. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/44235006

Asmolov, Gregory (2018). The Disconnected power of disinformation campaigns. *Journal of International Affairs*, vol. 71, no. 1.5, 2018, pp. 69–76. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26508120

Balkin, J. (2020). How to Regulate (and Not Regulate) Social Media. *Essays and Scholarship. Knight First Amendment Institute at Columbia University*. Retrieved 29 May 2021 from How to Regulate (and Not Regulate) Social Media | Knight First Amendment Institute (knightcolumbia.org).

Callahan, G. (2021). What are the potential risks of AI? Jan 5, 2021. *Blog. Artificial Intelligence*. Retrieved 29 May 2021, from https://www.rev.com/blog/what-are-the-potential-risks-of-artificial-intelligence.

GAO, United States Government Accountability Office (2018). Report to the Committee on Science, Space, and Technology, House of Representatives. Technology Assessment. Artificial Intelligence. Emerging Opportunities, Challenges, and Implications. Highlights of a Forum convened by the Comptroller General of the United States. March 2018, GAO -18-142SP. GAO-18-142SP, ARTIFICIAL INTELLIGENCE: Emerging Opportunities, Challenges, and Implications

Hunt, R., and F. McKelvey (2019). Algorithmic Regulation in Media and Cultural Policy: A Framework to Evaluate Barriers to Accountability. *Journal of Information Policy*, vol. 9, 2019, pp. 307–335. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/10.5325/jinfopoli.9.2019.0307

Kirby, R., K. Valaskova, J. Kolenick, and P. Kubala (2018). Online habits of the fake news audience: the vulnerabilities of internet users to manipulations by malevolent participants. *Geopolitics, History, and International Relations*, vol. 10, no. 2, 2018, pp. 44–50. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26802338

Krishna, D., N. Albinson, and Y. Chu (2017). Managing Algorithmic risks. Safeguarding the use of complex algorithms and machine learning. Deloitte report. Retrieved 3 June 2021, from Managing algorithmic risks (deloitte.com).

Levy, R. (2017). Taking Aim at Biased Algorithms. *Math Horizons*, vol. 25, no. 1, 2017, pp. 5–7. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/10.4169/mathhorizons.25.1.5

Metaxas, P. T., and E. Mustafaraj (2012). Social Media and the Elections. *Science*, vol. 338, no. 6106, 2012, pp. 472–473, Retrieved 29 May 2021 from www.jstor.org/stable/41703780

Morozov, E. (2017). Opposing the Exceptionalism of the Algorithm. *The Datafied Society: Studying Culture through Data*. edited by Mirko Tobias Schäfer and Karin Van Es, Amsterdam University Press, Amsterdam, 2017, pp. 245–248. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/j.ctt1v2xsqn.23

Prier, J. (2017). Commanding the Trend: Social Media as Information Warfare. *Strategic Studies Quarterly*, vol. 11, no. 4, 2017, pp. 50–85. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/26271634.

Reed C. (2018). How should we regulate artificial intelligence? Philosophical Transactions. Royal Society Publishing. A 376: 20170360. Retrieved 29 May 2021, from http://dx.doi.org/10.1098/rsta.2017.0360

Roese, V. (2018). You Won't Believe How Co-Dependent They Are: Or: Media Hype and the Interaction of News Media, Social Media, and the User. *From Media Hype to Twitter Storm,* edited by Peter Vasterman, Amsterdam University Press, Amsterdam, 2018, pp. 313–332. *JSTOR*. Retrieved 29 May 2021, from www.jstor.org/stable/j.ctt21215m0.19

Smith, A. (2020). Using Artificial Intelligence and Algorithms. Federal Trade Commission (FTC) Bureau of Consumer Protection. April 8, 2020, from Using Artificial Intelligence and Algorithms | Federal Trade Commission (ftc.gov).

Tutt, Andrew, An FDA for Algorithms (March 15, 2016). 69 Admin. L. Rev. 83 (2017), Available at SSRN: https://ssrn.com/abstract=2747994 or http://dx.doi.org/10.2139/ssrn.2747994.

Vasu, Norman, et al. (2018). Dissemination Techniques in Disinformation Campaigns. *S. Rajaratnam School of International Studies*, 2018, pp. 9–14, Fake news: national security in the post-truth era. Retrieved 29 May 2021, from www.jstor.org/stable/resrep17648.6

Vasu, N., B. Ang, T. Teo, S. Jayakumar, M. Faizal and J. Ahuja (2018). International Responses to Fake News. *S. Rajaratnam School of International Studies*, 2018, pp. 18–25, Fake news: national security in the post-truth era. Retrieved 29 May 2021, from www.jstor.org/stable/resrep17648.8

Vasu, N., B. Ang, T. Teo, S. Jayakumar, M. Faizal and J. Ahuja (2018). Unpacking Fake News. *S. Rajaratnam School of International Studies*, 2018, pp. 5–9, Fake news: national security in the post-truth era. Retrieved 29 May 2021, from www.jstor.org/stable/resrep17648.5

Weismann, S. (2019). How Not to Regulate Social Media. The New Atlantis No.58 (Spring 2019), pp. 58-64. Retrieved from How Not to Regulate Social Media on JSTOR.