REGULATING FACEBOOK

An ethical analysis of AI run amok, political posturing, and the failure of self-regulation

Major Research Paper

Submitted to Prof. Luc Juillet, GSPIA, University of Ottawa

Christina Lawand
Clawa022@uottawa.ca
Student ID: 0300032917

Table of Contents

ABSTRACT

The burgeoning field of AI ethics posits that AI is not just another utility that can be regulated once matured, but a powerful new form of smart agency that must be steered from the onset to benefit society.[1] But what happens when the proverbial train has already left the station? What are the challenges of regulating an AI application after its transformative power has been fully unleashed on society?

The purpose of this paper is to explore the challenges with ex parte regulation of AI, using one of its most notorious examples, Facebook, as a detailed case study. It demonstrates that regulation cannot be left up to the free market, and that policy makers must be more proactive in setting the parameters for the ethical development of artificial intelligence. While much has been written about the negative externalities of Facebook and other social media platforms, this paper offers a unique perspective by using the prism of AI ethics to evaluate how American political actors have responded to the Facebook controversy and to assess the company's efforts to self-regulate.

The paper uses a seminal moment in the company's history to explore ethical questions around the regulation of Facebook algorithms: CEO Mark Zuckerberg's appearance in extraordinary full day hearings before the U.S. Senate and House of Representatives in April 2018 to answer for his company's role in the Cambridge Analytica and Russian election interference scandals. Using five guiding principles (beneficence, non-maleficence, autonomy, justice and explicability) from the field of AI ethics as an analytical lens, a qualitative textual analysis of the legislative hearings was conducted to examine how issues were framed by Zuckerberg and U.S. legislators, how they were addressed prospectively, and what issues were missed from a public policy perspective. Policy considerations are further examined by evaluating Facebook's efforts to self-regulate in response to the hearings, and exploring potential pathways for regulators.

The paper concludes that while legislators were able to identify ethical challenges with Facebook algorithms in alignment with the five principles, effective regulatory proposals were hampered by a polarized congress, conflicting interpretations of the issues along partisan lines and a shallow understanding of the technology. It also shows that while self-regulation allows for more immediate corrective action, efforts taken by Facebook to date have been inconsistent, insufficient, and motivated by profit over ethics. Finally, this paper makes the case for better ex ante regulation of new AI technology that is not subject to the whims of political parties, with the need for clear legislative and regulatory boundaries, and independent oversight to protect the health of individuals, society, and democracy. In an era of rapid technological change that is expected to lead to ever greater societal disruptions, the Facebook story serves as an important cautionary tale on the need for a more proactive approach to policy-making.

[1] Floridi et al, (2018), p. 698.

INTRODUCTION

Artificial intelligence (AI) has the power to dramatically alter the way we work, travel and interact with each other. The burgeoning field of AI ethics posits that AI is not just another utility that can be regulated once matured, but a powerful new form of smart agency that must be steered from the onset to benefit society.[2] But what happens when the proverbial train has already left the station? What are the challenges of regulating an AI application after its transformative power has been fully unleashed on society?

The purpose of this paper is to explore the challenges with ex parte regulation of AI, using one of its most notorious examples, Facebook. It will examine how policy makers have struggled to fully comprehend the negative externalities of social media algorithms, and why Facebook's efforts to self-regulate have proven to be largely ineffective in reducing societal harm.

The story of Facebook provides an important case study of what happens with uncontrolled AI. It demonstrates that regulation cannot be left up to the free market, and that policy makers must be more proactive in setting the parameters for the ethical development of artificial intelligence. While much has been written about the negative externalities of Facebook and other social media platforms[3], this paper offers a unique perspective by using the prism of AI ethics to evaluate how American political actors have responded to the Facebook controversy and to assess the company's efforts to self-regulate.

---

[2] Floridi et al (2018), p. 689, with similar statements on the importance of planning and managing the development of AI in *Asimolar Principles* (2017) (principle #20) and on the need to guide ethical development as a key objective of the *Montréal Declaration for Responsible Development of Artificial Intelligence* (2018), for example.

[3] For example: Diebert, R (2019), Vaidhyanathan, S. (2018), Howard, P. (2020), Lafrance, A (2020), Zuboff (2019, 2021)

The paper uses a seminal moment in the company's history to explore ethical questions around the regulation of Facebook algorithms: CEO Mark Zuckerberg's appearance in extraordinary full day hearings before the U.S. Senate and House of Representatives in April 2018. At the time, Zuckerberg was summoned to answer questions about two major scandals involving his company: the Cambridge Analytica scandal involving the inappropriate use of data from millions of Facebook users, and Russia's abuse of the social network to interfere in the 2016 U.S. elections.

This paper will use the exceptional window provided by these hearings to empirically observe how issues were framed by Zuckerberg and legislators, how they were addressed prospectively, and what areas were missed from a public policy perspective. Notably, it will use five principles from the AI4 People's Ethical Framework for a Good AI Society[4] (beneficence, non-maleficence, autonomy, justice and explicability) as an analytical lens to examine:

- What ethical concerns with social media algorithms were identified by U.S. legislators, and what issues were missed;
- How perception and framing of the issues was divided along partisan lines, and how Zuckerberg's rhetoric differed from corporate practices;
- Why steps taken by Facebook to self-regulate since 2018 were insufficient to protect user privacy and to control the spread of misinformation and hate on the platform.

The paper will show that while legislators were able to identify ethical challenges with Facebook algorithms in alignment with the five principles, effective regulatory proposals were hampered by a polarized congress, conflicting interpretations of the issues along partisan lines and a shallow understanding of the technology. It will also show that while self-regulation allows for more immediate corrective action, efforts taken by Facebook to date have been inconsistent,

[4] Floridi et al, (2018)

insufficient, and motivated by profit over ethics. Finally, this paper makes the case for better ex ante regulation of new technology, with the need for clear legislative and regulatory boundaries, and independent oversight to protect the health of individuals, society, and democracy.

To set the contextual stage, this paper will begin with a background section on the Facebook controversy, including a brief history of Facebook regulation, and the company's use of intelligent algorithms as a central business feature. The methods section will then describe how the qualitative analysis was conducted, and provide an overview the AI4People's Ethical Framework for a Good AI Society as the main tool for analysis. An introduction to key concepts around artificial intelligence and AI ethics will help situate the framework within emerging literature in the field. In the third section, the five guiding principles from the framework will be used to analyze the discourse of legislators and responses from Zuckerberg during the extraordinary 2018 congressional hearings, including a discussion of what key themes emerged, how the discourse was divided along partisan lines, and what issues were missed by legislators. The fourth section will examine policy implications for regulators by observing how Facebook responded to ethical concerns about its technology, and how these actions fell short to the detriment of individual and societal health. This section will also highlight how the most recent U.S. regulatory actions against Facebook have proven to be insufficient, and propose potential pathways for regulators in the U.S. and abroad to fix social media and ensure the ethical development of algorithms. Finally, the conclusion will speak to the broader implications of the Facebook controversy, and how it serves as an important cautionary tale on the need for "ex ante" regulation in an era of rapid technological change.

1. BACKGROUND ON THE FACEBOOK CONTROVERSY

1.1. "Break things and move on": a brief history of Facebook regulation

The founding story of Facebook, and its rise to the pinnacle of corporate success, is now legendary. Started by Mark Zuckerberg in his Harvard dorm room as an online platform to rate campus women, Facebook has morphed into a global behemoth, valued at more than $780 billion, with 2.5 billion users around the planet.[5] It has also revolutionized the way people connect with each other, the way businesses advertise to consumers, and how political campaigns are conducted.

Like other social media companies and much of the internet, Facebook has operated in a largely unregulated environment. Between 2004 and 2014, the company's motto was to "break things and move on", a clear signal of Zuckerberg's drive to push the boundaries of technology and disrupt the information ecosystem. The company's history is marked by a pattern of rapid growth, technological innovation, and exploitation of its business model to maximize profit – followed by an ethical breach or controversy involving the application of its technology – followed by an apology from Mark Zuckerberg, and a promise to "do better".[6]

From a regulatory perspective in North America, Facebook was first taken to task in Canada, when the Privacy Commissioner's Office ruled in 2009 that the company did not have adequate safeguards in place to prevent unauthorized access of data and failed to ensure meaningful consent was obtained by users of the social media platform.[7] In 2011, the U.S.

---

[5] Current market valuation in U.S. dollars at: https://finance.yahoo.com/quote/FB/key-statistics/

[6] In *U.S House of Representatives* (April 2018) pp. 39-40: Rep. Schakowsky quoted all the times Zuckerberg apologized, starting at Harvard.in 2003: "'I apologize for any harm I have done as a result of my neglect." 2006. "We really messed this one up." 2007. "We simply did a bad job. I apologize for it." 2010. "Sometimes we move too fast." 2011. "I am the first to admit that we have made a bunch of mistakes."

[7] Office of the Privacy Commissioner of Canada (Aug., 2009)

Federal Trade Commission charged that Facebook made "unfair and deceptive claims" to consumers about their privacy, and issued an order requiring the company to better inform users about how their personal data was being collected and used, and to perform regular privacy audits.[8]

Facebook promised modifications to its platform to comply with the rulings, but two seminal moments would later bring the company under extraordinary legislative scrutiny in 2018: the Cambridge Analytica scandal, involving the inappropriate sale of data from millions of Facebook users to political consulting firms, and the U.S. Department of Justice investigation on Russian election interference. A February 2018 indictment accused Russian operatives from the Kremlin-backed Internet Research Agency of using false personas to purchase ads and spread misinformation on Facebook against Hillary Clinton in order to help elect Donald Trump in the 2016 election.[9]

Under public pressure, Zuckerberg agreed to take a "broader view" of Facebook's responsibilities and propose changes to its platform, including imposing new restrictions for advertisers and third-party applications, hiring an army of 20,000 fact-checkers, and offering more transparency for users and more consent options over how their data is used.[10] In May 2018, the European Union's General Data Protection Regulation (GDPR) imposed strict new data privacy and transparency requirements for all companies doing business in Europe, forcing Facebook to improve privacy protection.

---

[8] U.S. Federal Trade Commission (Nov. 2011): https://www.ftc.gov/news-events/press-releases/2011/11/facebook-settles-ftc-charges-it-deceived-consumers-failing-keep
[9] United States Justice Department (Feb. 16, 2018)
[10] From Zuckerberg deposition to United States Senate (April 10, 2018):

However, privacy is merely one of the problematic features of Facebook and other algorithm-driven social media platforms. Zuckerberg and the CEOs of Twitter and Google were summoned to appear before the U.S. Senate on two separate occasions during the hotly contested 2020 election race, a testament to the perceived power of their platforms in influencing the election outcome. The big tech CEOs faced attacks from both sides of the aisle, but for reasons clearly split along ideological lines[11]. Ironically, the big tech companies were caught in the maelstrom of the polarization that they unwittingly helped to unleash.

Facebook is now facing its largest regulatory hurdle to date, but the action is grounded in economic rather than moral principles. In December 2020, the U.S. Federal Trade Commission accused the company of anti-competitive practices and Facebook was hit with an antitrust lawsuit, supported by 48 states.[12] The company is also targeted in new initiatives to regulate the internet in countries such as the U.K. and Australia,[13] and is facing fresh public and legislative scrutiny in the U.S. for its role in spreading misinformation that led to COVID-19 science denial and the storming of the Capitol.

1.2. Background on Facebook's use of algorithms and their negative externalities

Artificial intelligence has played a central role in Facebook's astounding growth. Three key design features of the social media platform include: a) a business model based on harvesting large amounts of data from individual users, known as "surveillance capitalism"[14];

[11] New York Times live stream (Oct. 28, 2020) https://www.nytimes.com/live/2020/10/28/technology/tech-hearing
[12] Federal Trade Commission (Dec. 9, 2020)
[13] Ofcom (Dec. 15, 2020) and Packham, C. (Feb. 11, 2021)
[14] The term "surveillance capitalism" was coined by Harvard professor Shushana Zuboff in her book on the topic, Zuboff, S. (2019)

b) AI-driven algorithms that help personalize the Facebook user experience; and c) technical design features that encourage users to spend as much time as possible on the platform.

These features are problematic for a number of reasons according to recent studies. First, people are being surveilled without their knowledge; Facebook's consent forms have historically been opaque and made it difficult for consumers to understand what they are agreeing to and how much of their data is being shared.[15] Applications offered on Facebook, such as a quizzes or games, double "as a means to observe you and acquire data about you."[16] Harvesting vast amounts of data has allowed Facebook to sell ads by customizing content and targeting individuals based on their preferences and behaviours, a practice known as "microtargeting"[17]. The efficacy of microtargeting to sell things or ideas to people, enabled by vast quantities of personal data and sophisticated algorithms, have completely disrupted traditional advertising, the main source of revenue for mainstream media. By providing free content and gobbling up the bulk of ad revenues, big tech companies such as Facebook and Google have broken the business models and caused the demise of many local newspapers and independent news organizations across North America and around the world.[18]

Second, Facebook strives to give users "the most personalized experience" with algorithms that distribute, organize and rank content in their news feeds, based on what individuals like and their online behaviour. This leads to what Cass Sunstein has described as the

[15] Diebert, R. (2019) p. 28
[16] Ibid
[17] For Philip Howard, director of the Oxford Internet Institute, contemporary microtargeting "involves preparing and delivering a message that is customized for particular individuals using their data, social ties, cognitive biases, and big data records, often in ways that are unsupervised, untraceable, unaccountable and unknown to the individuals." In Howard, P. (2020) – p. 174
[18] Grieco, E. (2020); OECD (2010)

echo chamber effect,[19] where people are only exposed to the views and ideas that they already share, and that are reinforced in their "filter bubble" – the term coined by Eli Paliser to describe the unique, personal cocoon of information created by algorithmic filter technology on social media and the internet.[20] The ubiquity of Facebook and social media in modern society, with its echo chamber and filter bubble effects, leads to people living in alternate realities and the polarization of political views.

Third, Facebook leverages research in behavioural psychology to make its algorithms more profitable. (It's not a coincidence that Zuckerberg majored in both computer science and psychology). This includes technical features designed to keep people engaged with the content and maximize the amount of time they spend online. For example, according to Facebook's first president, Sean Parker, the "like" button was designed to create a constant feedback loop and give users a "little dopamine hit" – or produce chemicals in the brain associated with addiction.[21] Algorithms ensure that users are continuously fed a stream of information that they want to interact with. As Parker explained, Facebook's objective was to "consume as much of your time and conscious attention as possible".[22]

Finally, due to the vast amount of information now available online, people's attention has become a scarce commodity. This has led to what Nobel prize-winning economist Herbert A. Simon has coined as "the attention economy", positing that "a wealth of information creates a poverty of attention."[23] With anyone being able to share their views with the rest of the world

[19] Sunstein, C. (2007), p. 7 and p. 122: Sunstein first introduced the echo chamber hypothesis in 2002 as a possible outcome of the internet age and elaborated on it in 2 subsequent books in 2007 and 2017, based on the flourishing of personal blogs (2007) personalized search engines, and social media (2017).
[20] Paliser,E. (2012)
[21] Solon O (2017)
[22] Ibid
[23] As quoted in Mintzer, A. (March 31, 2020)

anywhere and at any time, it is often the most outrageous views that cut through the noise and solicit a reaction from social media users. This in turn has led algorithms to promote provocative or attention-grabbing information in people's news feeds in order to encourage people to spend more time online and by extension, be exposed to more ads.[24]

The prioritization of provocative over truthful information – coupled with the decline of mainstream media as an information "gatekeeper" – has led to the proliferation of fake news and conspiracy theories on Facebook and other social media platforms. The immediacy of social media means that lies can quickly go viral, and studies have found that efforts to quash rumours through direct refutation often leads to further propagation and acceptance.[25] These inherent features of Facebook algorithmic design have not only contributed to the rapid diffusion of lies, they have fueled cynicism, amplified extremism, and been handily exploited by illiberal populists and authoritarians alike.[26] Facebook advertising and microtargeting services have been used to spread misinformation in the Brexit campaign, to suppress minority voters in the U.S., and to interfere in elections.[27] Algorithms were also exploited organically, to organize the January 6 protest and siege of Washington's Capitol building at the behest of a post-truth President. The weaponization of lies, via the pulpit of social media, is a central feature of the Kremlin's hybrid warfare tactics, both to crush opposition voices close to home and as part of Russia's broader geo-political power ploy to destabilize the Western block.[28]

[24] Silverman, C (2015)
[25] Berinsky, A. J. (2017) p. 241
[26] See for example in Diebert, R. pp. 34-35
[27] Howard, P. (2020) pp. 38-74
[28] Ibid and Barrett, P. et al (2018) for example.

2. METHODS AND FRAMEWORK

The trouble with Facebook is due to uncontrolled development in an era of exponentially rapid technological transformation. AI has enormous potential for good, but can also lead to negative externalities if not developed thoughtfully. To understand the ethical shortcomings of Facebook's AI-driven business model and policy deliberations around them, this paper is using the AI4People's "*Ethical Framework for a Good AI Society*" as the main lens for analysis.[29]

U.S. congressional committee hearings provide a rich body of empirical evidence to analyze legislative discourse from an ethical perspective. Zuckerberg has made a total of five appearances before the U.S. Senate and the U.S. House of Representatives over the past three years alone. This paper will focus primarily on the April 2018 U.S. hearings held in the wake of the Cambridge Analytica and Russian election interference scandals for the following reasons:

- The extraordinary nature of the hearings at a watershed moment for the company; Zuckerberg appeared in two full days of hearings in both chambers of Congress
- The 2018 time period allows for both a retrospective review of Facebook's actions that led up to the scandals, and allows for prospective scrutiny of promises made by Facebook at the time to improve its technological tools and processes
- The rich body of material submitted for the evidentiary record include the hearings themselves, third-party depositions, and hundreds of follow-up written questions from legislators subsequently answered by Facebook[30]

[29] Floridi et al (2018). AI4People is a mutli-stakeholder forum that brings together "all actors interested in shaping the social impact of new applications of AI". The forum was spearheaded by Oxford University's Digital Ethics Centre and the Atomium European Institute for Science, Media and Technology.
[30] They are also reliable statements of record as perjury laws apply to the evidence of witnesses provided before congressional committees

- In his home country, Zuckerberg agreed to appear before the congressional committee hearings in person, unlike the U.K. parliamentary hearings on the Cambridge Analytica scandal; a British parliamentary committee held Zuckerberg in contempt for sending junior representatives as witnesses who were unable to answer many questions.[31]
- The oversized influence of U.S. politics and policies on global stability and a Liberal world order; the hearings therefore provide a microcosm of the influence of technology on larger domestic and international geo-political dynamics.

The AI4People ethical framework was used to classify and organize legislative discourse at the 2018 hearings according to ethical principles. A qualitative analysis of the oral statements and questions from 51 U.S. senators (representing more than half of the Senate) and 56 members of the House of Representatives was conducted from 251 pages of oral hearing transcripts, including Zuckerberg's responses. The discursive arguments were then labelled and categorized according to the framework's five guiding ethical principles - namely beneficence, non-maleficence, autonomy, justice and explicability – and their sub-categories (for example, "explicability" includes the concepts of both "transparency" and "accountability" which were separately labelled). Oral arguments reflect how political and corporate actors perceive the issues, and how they want to frame them for public consumption. The analysis also examined written follow-up questions and answers submitted by 30 senators to Zuckerberg as part of the 2018 hearings. While more technical in nature, the 218 additional pages of transcripts offered useful insights into what issues legislators were seeking to better understand from an ethical

[31] House of Commons Digital, Culture, Media and Sport Committee. (February 2019) Zuckerberg was also threatened with being held in contempt for refusing to appear at a hearing for international legislators on "Big Data, Privacy and Democracy", held in Ottawa in May 2019. See: https://www.ourcommons.ca/DocumentViewer/en/42-1/ETHI/news-release/10470470

perspective about Facebook's business model, what regulatory avenues they were trying to explore, as well as what lines of inquiry were missed altogether in the legislative probe. The follow-up questions and answers also helped to clarify Facebook's official corporate policy, and distinguish it from some of Zuckerberg's more aspirational rhetoric at the committee hearings.

Hearing transcripts were supplemented by government and media reports to understand actions taken by the company and regulatory authorities since 2018, as well as to document more recent controversies involving Facebook during the global pandemic and 2020 U.S. election campaign. The paper also uses academic and grey literature to test the assumptions of legislators, evaluate Zuckerberg's claims, as well as to contextualize and further elaborate on areas of ethical concern with social media algorithms and their impact on individual and societal health. These literature were also used to explore potential pathways for policy makers to regulate social media specifically, and AI more broadly. Finally, Ulbrecth Beck's "Risk Society" theory of modernization will be referenced to discuss lessons learned from the Facebook experience, by illustrating that the technology built to improve one aspect of our lives may be putting us at risk by transforming many other aspects of our lives.

2.1. AI Ethics: Key concepts and a framework for analysis

What is AI?

Facebook is a pioneer in the field of artificial intelligence (AI) and a leading global player in the AI "arms race" along with other big tech companies such as Google, Apple and Amazon. From an epistemological perspective, AI is difficult to define, because the boundaries of what is considered to be "intelligent" as a machine function is constantly evolving.[32] (For

[32] Scherer, M. (2015) pp. 354.

example, in the past, one could argue that calculators were "intelligent" as they could solve equations in a fraction of the time of a human being).

AI pioneers Stuart Russell and Peter Norvig offer this widely referenced definition: "Systems that think like humans, systems that act like humans, systems that think rationally, systems that act rationally."[33] While the technology may not yet have reached human equivalency, AI is now broadly defined as computer systems able to perform tasks that "normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."[34] For the purposes of this paper, the focus is on systems that can make judgments or decisions of social, economic or physical consequences that were previously made by humans. The trifecta of factors that have enabled AI technology to take flight over the past 15 years are: 1) the ability to collect and store large quantities of data (known as the "big data" era) 2) greater network computational power and 3) algorithms – or the complex equations used to analyze the data and make automated decisions.[35] With machine learning, algorithms are able to learn from their mistakes and modify their decisions without human intervention.

About AI Ethics

AI has enormous potential for good, but also comes with great risk if not developed thoughtfully. The Montreal Declaration for a Responsible AI frames it this way: "*From an ethical point of view, the development of AI poses previously unknown challenges. For the first time in history we have the opportunity to create non-human, autonomous and intelligent agents*

[33] Russel, S. and Norvig, P. (1995)
[34] Dubber, A.M. et al (2020)
[35] Scherer, M. pp. 355-358

that do not need their creators to accomplish tasks that were previously reserved for the human mind. These intelligent machines do not merely calculate better than human beings, they also look for, process and disseminate information."[36]

The idea of "robot ethics" predates moral discussions of artificial intelligence writ large[37], but unlike social media algorithms, robots are visible, and humans are aware when they are interacting with them. Facebook algorithms work in the shadows to sift through vast quantities of data and determine what individual users get to see and interact with on their platforms. Beck's risk society theory posits that we often cannot see the technology that is putting us most at risk. And while Beck was discussing the impacts of radiation at Chernobyl as an example of technological risk with no obvious physical presence, the same logic applies to social media algorithms, with its discrete but profound effects on individuals and society.[38]

AI Ethics is a relatively new, but rapidly expanding field; most of the academic literature can be found in the past five years, or a good decade after the creation of Facebook. Leading think tanks embody a multidisciplinary approach, involving the fields of law, applied sciences, computer science, and social sciences [39]– or a multistakeholder approach including industry and/or government. "The Partnership on AI" is one of the earlier and better known initiatives launched by the big tech companies, including Apple, Facebook, Google and IBM, to brand themselves as responsible stewards of new technology. [40]

---

[36] Montréal Declaration for the Responsible Development of AI. (2018)
[37] Earliest example hails from science fiction; Isaac Asimov's "Robot Laws" (1942) is widely referenced in literature about AI regulation and ethics.
[38] Beck, U. (1992)
[39] Examples include AI Now, (NYU) Oxford University's Digital Centre of Ethics
[40] See Partnership on AI. https://www.partnershiponai.org/

The central thesis of this paper is on the need for more ethical considerations upfront to mitigate the negative externalities of artificial intelligence, using Facebook as a case study. The AI4 People's "*Ethical Framework for a Good AI Society*" is a widely referenced document that provides a useful lens for analysis. The authors start from the premise that AI is not inherently bad, and that misplaced fear or excessive reaction can present opportunity costs, by underusing the new technologies to the detriment of society. In this sense, it offers a counterweight to emerging critical literature focusing primarily on the dangers of artificial intelligence. However, the authors also concede that AI is not just another utility, that can be regulated once matured, but a powerful new form of "smart agency" that must be steered from the outset to benefit the greater good, and to avoid the pitfalls of misuse or overuse.[41]

The framework describes four opportunities and corresponding risks associated with the use of AI.[42] These include:

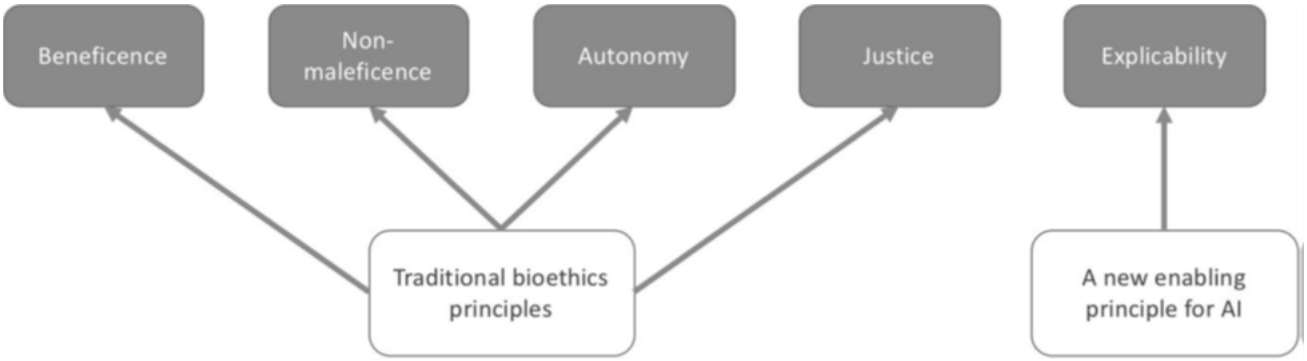
1) *Enabling human realization without devaluing human abilities,* AI can allow humans to spend more time on rewarding activities by automating mundane tasks; the corresponding risk is that humans will lose their skills in sensitive domains, creating vulnerabilities in the case of AI malfunction.
2) *Enhancing human agency without removing human responsibility*: if developed thoughtfully, AI can multiply the ability of humans to understand complex problems and act on them, (for example sequencing the human genome to help develop the coronavirus vaccine); on the flip side, delegating important decisions to machines can erode human responsibility.

[41] Floridi et al, p.689
[42] Ibid, pp. 690-694

3) *Increasing societal capabilities without reducing human control*: from optimizing resources to helping with the prevention and cure of disease, AI has the power to radically enhance what we are capable of as a society; however an overreliance can lead to a loss of human control over important decisions that should remain subject to human supervision and choice.
4) *Cultivating societal cohesion without eroding human self-determination*: AI solutions can contribute to a more collaborative society, provided that all stakeholders are involved in their design and accept their intended use. However, predictive power can come at the expense of human dignity if it does not consider or allow for individual choice.

To further consider ethical concerns and classify policy deliberations around social media algorithms, five over-arching principles will be examined from the framework.

Figure 1



Source: *Floridi et al, p. 700.*

These principles were distilled from key international initiatives on the responsible use of artificial intelligence, including the Montréal Declaration[43] and the Asilomar Conference[44]. They

[43] Montréal Declaration for the Responsible Development of AI
[44] Asilomar AI Principles. (2017)

include four principles borrowed from the field of bio-ethics, and one new principle specific to AI: the principle of explicability. The five principles are synthesized below: [45]

- **Beneficence**: to ensure that AI technologies benefit societies, promote well-being, and empower as many people as possible. This includes protecting human dignity at an individual level, and promoting sustainability at a planetary level, by ensuring technology is used to help people prosper and to preserve a "good environment for future generations".
- **Non-maleficence**: this includes not just preventing harm from the nefarious use of AI, but ensuring that smart technologies operate within secure constraints and protect the privacy of individuals to avoid "unintentional" harm, or create "new harm" by undermining existing social structures. The principle aims to prevent *overuse* of AI, which can cause accidental harm to individuals and society, as well as *misuse* of AI, or the deliberate harm of others by nefarious political or criminal actors.
- **Autonomy**: the idea that humans should have a choice over what decision-making power they get to keep for themselves. This speaks to the question of human control. The added complication with AI is that unlike in bioethics, where we can clearly delineate the boundaries of human decision-making (for example, in offering choices for treatments), with smart agency we *willingly* cede decision-making power to machines. The central question then becomes what types of decisions human want to keep for themselves rather than delegate to machines.
- **Justice:** The AI ecosystem is "inherently unequal", with a small fraction of humanity engaged in the design and development of technologies that are transforming the lives of

[45] Floridi, L. et al, pp. 696-700

everybody else.[46] The objective of this principle is to ensure the benefits of technology are broadly distributed. From a human rights perspective, AI should aim to promote equity, seek to eliminate discrimination, and share prosperity.

- **Explicability:** both in the epistemological sense of "intelligibility" (AI decisions must be transparent and explainable) and in the ethical sense of "accountability", or the ability to contest AI decisions and to hold human agents responsible for them. This principle complements the other four principles, because to determine if a form of smart agency is beneficient or non-maleficent, if it provides humans with sufficient autonomy, or if potential benefits are equally distributed, the first step is to understand how the technology actually works, and – in case of failure – who is responsible for it.

The principle of "explicability" has been the focus of much emerging critical thought about artificial intelligence. In particular, concerns have been expressed about the so-called "black box" of machine learning, or the idea that questions can be answered or decisions can be made by machines that can affect people's lives, without the ability to understand how those decisions were arrived at, or on what basis they could be contested.[47] This is the dimension that uniquely sets AI apart from other forms of technologies, biological or otherwise.

The Ethics of AI Ethics

Critiques of current approaches to AI Ethics, including the proliferation of frameworks and guidelines, center on both conceptual ambiguities, and a lack of implementation or enforcement. A systematic mapping of AI guidelines internationally found that while global convergence is emerging around five ethical principles (transparency, justice and fairness, non-

[46] Floridi et al, p. 699
[47] Examples include : Campolo, A. et al (2017); Pasquale, F. (2015); O'Neill, C. (2015)

maleficence, responsibility and privacy), there remain "substantive divergence in relation to how these principles are interpreted, why they are deemed important, what issue, domain or actors they pertain to, and how they should be implemented."[48]

Research in ethics and in AI, respectively, involve distinct scholarly communities, leading to terminological ambiguities around key concepts such as "agency, autonomy, intention, and responsibility", as well as confusion over how to apply these concepts to machine-made decisions.[49] In addition, both ethicists and engineers can have trouble anticipating the downstream effects of technology. This can sometimes lead to important omissions. For example, a systematic review of 21 prominent AI Ethics guidelines (including the AI4 People Framework) found that virtually none of these documents dealt explicitly with the "hidden" social and ecological costs of AI systems, such as the reduction of social cohesion by AI ranking and filtering on social networking sites.[50] Indeed, until recently, some of the most prominent international initiatives on AI ethics focused on future or specialized applications of smart technology (such as autonomous weapons or predictive policing), while completely ignoring the more ubiquitious use of algorithms by social media companies.[51]

Another key challenge of AI guidelines is their lack of teeth. Critics argue that principles are vague, elastic in scope, and currently operate as an "empty vessel into which anyone (including the tech industry and the so-called Digital Titans) can pour their preferred ethical content."[52] With little external regulation of AI development, ethics codes are often adopted by

[48] Jobin, A., Ienca, M. & Vayena, E. (2019) p. 389
[49] Powers, T. and Ganascia, J.G. (2020)
[50] Hagendorff, T. (2020). p.105
[51] For example, the Asimolar conference (2017) and the international forum and workshops that led to the Montréal Declaration (2018) did not consider social media algorithms in their deliberations.
[52] Slee (2020)

industry to show themselves as responsible corporate citizens. However, enforcement is problematic with a soft law approach to regulation, as deviation from industry codes often have no clear consequences. For this reason, big tech companies embrace ethics within their institutions primarily as a "marketing strategy" and a tool of public appeasement, in order to avoid more stringent constraints on AI development codified in law.[53] However, when it comes to public regulation of industry, ethical considerations often take a back seat to economic interests.[54]

3. THE FACEBOOK LEGISLATIVE HEARINGS: A QUALITATIVE ANALYSIS

The 2018 Facebook hearings were described as "extraordinary" and a "watershed moment" in the history of social media regulation.[55] On April 10, Zuckerberg testified alone before a joint committee representing half of the U.S. Senate for a full day of hearings, and repeated the exercise the following day before the U.S. House of Representatives.

In his opening remarks and deposition, Zuckerberg addressed the scandals that had brought him here. In 2013, Cambridge University researcher Alexsander Kogan developed a personality quiz application that was hosted on the Facebook platform. More than 300,000 users signed up and agreed (whether wittingly or not), to share some of their Facebook information, as well as the information of their friends, even if the friends in question had not signed up for the app. In 2015, a *Guardian* news article revealed that Kogan shared data collected from the app with Cambridge Analytica, a political consulting firm supported by key figures of the Brexit and Trump campaigns. In response to the article, Zuckerberg said Facebook had asked the firm to

[53] Hagendorff, p.113
[54] Harrington (1996) p. 363
[55] United States Senate (2018) p.3 and p.81

certify that it had deleted all data acquired without user consent. In March 2018, *The Guardian* and the *New York Times* reported that Cambridge Analytica may not have deleted the data as certified, exposing the data of 87 million Facebook users to political manipulation and prompting an investigation from the U.K. Information Commissioner.[56]

In his statement of facts about Russian election interference in 2016, Zuckerberg confirmed that Russia's Internet Research Agency (IRA) spent more than $100,000 on Facebook ads in the U.S. election campaign that reached an estimated 11 million U.S. voters. The IRA also launched an organic disinformation campaign, with 42 accounts that generated 80,000 Facebook posts, reaching an estimated 126 million people.[57]

This section of the paper examines a) what issues were identified at the hearings from an AI ethics perspective and how they were publicly framed, b) partisan differences over how to interpret the issues, and c) what issues were missed by legislators.

3.1. Ethical issues identified by legislators, according to the 5 principles

Zuckerberg reasonably faced a barrage of questions from both sides of the aisle, with the issue of how Facebook had compromised the public's trust a common bipartisan theme. An ethical analysis of the legislative discourse shows that all five principles of the AI4 People's framework were addressed to a varying degree throughout the two-day hearings – with most of the questions from members focusing on the principles of non-maleficence (related to the breach of user privacy and the platform's abuse by bad actors) and explicability, or the need for

[56] U.S Senate pp. 9-10 and U.S. House of Representatives (2018) p.11
[57] U.S. Senate pp. 12-13

Facebook to be more transparent and accountable. The main arguments are organized by principle below:

a). <u>Beneficence</u>: Arguments that Facebook's AI systems benefit society and promote well-being came primarily from Zuckerberg himself, though some G.O.P. House members praised the company as an American entrepreneurial success story, while others noted that the platform helped them to stay connected with their constituents. Due to the ubiquity of Facebook, legislators broadly acknowledged that it was simply not possible for them to NOT have a Facebook account.

Zuckerberg argued that societal well-being is central to Facebook's mission. "*Facebook is an idealistic and optimistic company (…) focused on all the good that connecting people can do. And as Facebook has grown, people everywhere have gotten a powerful new tool for staying connected to the people they love, for making their voices heard, and for building communities and businesses*."[58] The platform, he said, helped to organize the #metoo movement, raise funds for hurricane relief, and has become a virtual front office for more than 70 million small businesses.

Facebook's algorithms are beneficent, argues Zuckerberg, as they allow people to receive personalized content that is targeted to their interests and whereabouts. For example, if a Facebook user likes pages or belongs to online groups for downhill skiers, they will receive ads from the local ski shop – or if they are traveling, promotions to the nearby ski hill. And while most people say they don't like ads, Zuckerberg says what they really mean is that they don't like ads that are not relevant. "*The overwhelming feedback that we get from the community is*

[58] Idem – p. 8

that people would rather have us show them relevant content than not."[59] Facebook's surveillance-based business model also benefits society according to Zuckerberg, by providing a free service and allowing everyone to have a voice, therefore enabling the Habermasian ideal of advancing society through the free exchange of ideas. "*Offering an ad-supported service is the most aligned with our mission of trying to help connect everyone because we are trying to offer a free service that everyone can afford*."[60] A senator praised what he called Facebook's "*positive role"* in African countries with authoritarian regimes by providing a voice to opposition figures in places where the only mainstream media is state-run media.[61]

b). <u>Non-Maleficence</u>: Non-maleficence emerged as a common theme throughout the two days of hearings, with many questions centred on the lack of privacy protection and security for Facebook users, the abuse of the platform by malicious actors, and on the societal harm caused by undermining existing social structures. "*For all the benefits that Facebook has provided in building our communities and connecting families, I think a devil's bargain has been struck*". [62]

Members on both sides of the aisle accused Facebook of having a history of lax data practices and failing to adequately protect user privacy.: "*Facebook is just the latest in a never-ending string of companies that vacuum up our data but fail to keep it safe*." [63] Members of the Democratic party in particular questioned the company's extensive surveillance apparatus -"*from the moment we wake up in the morning until we go to bed at night, we are on tablets, phones and*

[59] Idem. p.9
[60] Idem p. 16
[61] Idem, Sen. Flake, J. p. 106.
[62] U.S. House of Representatives, Hon. Castor, K. p. 49
[63] Idem, Hon. Pallone, F. p. 5

being tracked"[64] – with Senator Richard Durbin rhetorically asking Zuckerberg if he would feel comfortable sharing the name of the D.C. hotel where he had spent the prior night. For Republican Committee Chair Chuck Grassley, the Cambridge Analytica scandal represented "*a breach of trust and likely improper transfer of data (…) igniting a larger discussion on consumers' expectations and the future of data privacy in our society*."[65] However, some Republican legislators were not as troubled by the company's privacy lapses. "*I am a proud member of Facebook*," explained Senator Thom Tillis. "(…) *if you don't want to share something with Facebook then don't share it."*[66]

Nefarious use of the platform by foreign agitators was an area of high concern for Democratic lawmakers, particularly in the wake of Russian election interference that favoured Donald Trump. House Committee Co-Chair Frank Pallone called Facebook an "*unwitting co-conspirator in Russian influence campaigns*", and accused the company of being caught flat-footed in 2016. "*How was it so blind to what the Russians and others were doing on its systems? Red flags were everywhere. Why didn't anyone see them or were they ignored*?"[67] The Democratic Senator from Wisconsin raised that swing states such as Wisconsin were a "*major target of Russia-bought ads*"[68], with the platform's tools used to target voters based on racist, anti-Muslim or anti-immigrant views. Zuckerberg admitted that "*one of my greatest regrets in running the company is that we were slow in identifying the Russian information operations in 2016*."[69] However, when asked in a written follow-up question by Senator Kamala Harris how much revenue Facebook had earned from foreign propaganda in the 2016 election, the company

[64] U.S. Senate, Sen. Nelson, B. p. 5
[65] Idem, Grassley, C. (Committee Chair) p. 5
[66] Idem, Sen. Tillis, T., p.28.
[67] U.S. House of Representatives, Hon. Pallone,F. p. 6
[68] U.S. Senate, Sen. Baldwin, T. p. 131
[69] U.S. Senate, Zuckerberg, M. p. 19

dodged the question, stating that "*revenue that is attributable to inauthentic or false accounts is immaterial*."[70]

Members highlighted other examples of malicious use of the platform, both domestically and internationally. Most egregiously, a UN investigation found Facebook to be "*a breeding ground for hate speech*" against Rohingya refugees and played a role in inciting genocide in Myanmar. Congressional members from the mid-West also pointed to the use of Facebook to sell illegal opioids online, fuelling a deadly national opioid crisis in the American heartland.

Finally, some concerns were expressed primarily from Democrats about the broader societal harm caused by the misuse and abuse of Facebook algorithms. "*I think the damage done to our democracy relative to Facebook and its platform being weaponized are incalculable."*[71] "*We have seen the scale and reach of extremism balloon in the last decade, partially because of the expansion of social platforms, whether it is a white supremacist rally in Charlottesville that turned violent, or to ethnic cleansing in Burma that turned into the second largest refugee crisis in the world*." [72] "*While securing our privacy is necessary it is not sufficient. We must take steps immediately to secure our democracy*.[73]" Arguments that elicited a mea culpa from Zuckerberg: "*It is clear that we did not do enough to prevent these tools from being used for harm as well and that goes for fake news" (…) I don't want anyone to use our tools to undermine democracy. That's not what we stand for."*[74]

[70] Idem, written follow-up question from Harris, K., p. 370
[71] U.S. House of Representatives, Hon. Eshoo, A. p. 23
[72] Idem Hon. Engel, E. p. 27
[73] Idem Hon. Pallone, F. p. 6
[74] U.S. Senate, Zuckerberg, M. p. 13

c). **Autonomy:** The issue of autonomy centred mostly on what control users have over their own data, rather than on what decisions should be controlled by machines on the social media platform. Zuckerberg argued that the question of data ownership was clear for his company. "*The first line of our terms of service say that you control and own the information that you put on Facebook*".[75] However, for many members, the turgidity of Facebook's terms of service agreements makes it difficult for users to exercise autonomy and make clear choices. "*Your user agreement sucks (…) The purpose of that user agreement is for Facebook to cover their rear end. It is not to inform your users about their rights. (…) As a Facebook user are you willing to give me more control over my data?*"[76] Others questioned Zuckerberg on how much ownership users actually have on what has become Facebook's most valuable commodity. "*I cannot imagine that it is true as a legal matter that I actually own my own Facebook data because you are the one monetizing it (…) otherwise we would be getting a cut*". [77]

d). Justice: Many members raised questions about whether the benefits of Facebook algorithms were equally distributed, and whether the platform treated people fairly across the political and demographic spectrum. However, the types of accusations of bias were clearly split along ideological lines. Leading the charge on the Republican side, Senator Ted Cruz accused Facebook and other big tech companies of "*a pervasive pattern of bias and political censorship*" against Conservative and Evangelical Christian voices (…) and of "*purposefully and routinely suppressing conservative stories from trending news*".[78] The Republican Senate Committee

[75] Idem. p. 16
[76] Idem, Sen. Kennedy, J. p. 127
[77] Idem, Sen. Schatz, B. p. 95
[78] Idem, Sen. Cruz, T. p. 49

Chair asked whether Facebook was tipping the scales against Conservatives from a technological perspective: "*Do you subjectively manipulate your algorithms to prioritize or censor speech*"? [79]

On the other side of the aisle, accusations of algorithmic bias centred on discrimination against minorities. "*Unfortunately, we have seen how (…) technology platforms like Facebook can actually be used to double down on discrimination, and give people more sophisticated tools with which to discriminate."*[80] Senator Cory Booker cited a lawsuit against Facebook for discriminatory practices in advertising, accusing the company's algorithms of "redlining" ads for housing, employment, credit, and public accommodation by limiting ad reach based on gender, ethnic affinity, and other socio-demographic characteristics. A lack of diversity in the tech sector was also identified as problematic, and contributing to bias against minorities in algorithm development.[81]

e). **Explicability:** Zuckerberg faced rebuke from both parties over Facebook's lack of transparency about how it uses data, and its lack of accountability over privacy breaches. Several members took issue with Facebook's complicated privacy settings and opaque terms of service, featuring 3200 words with 30 links, and a data policy with 2700 words and 22 links. **"***People have no earthly idea what they are signing up for.*"[82] Facebook's business model provides personal information in exchange for free services but "*for this model to persist, both sides of the bargain need to know the stakes that are involved (…) at a minimum consumers must have the*

[79] Idem, Sen. Grassley, C. p. 33
[80] Sen. Booker, C. p. 116
[81] Idem p. 117
[82] Idem. Sen. Schatz, B. p. 95

transparency necessary to make an informed decision about whether to share their data and how it can be used." [83]

From an accountability perspective, Zuckerberg was asked why it took 27 months to notify Facebook users affected by the misappropriate use of their data by Cambridge Analytica after it was first reported by the Guardian newspaper in 2015. Users were only contacted by Facebook after the public uproar over fresh reports of the scandal in 2018. "*Why were 87 million Facebook Users not notified that their personally identifiable information was taken without their consent,"* asked Democratic Senator Bill Nelson. *"Do you not feel you have an ethical obligation to notify (them)?*"[84] Others echoed the need for tech companies to take greater responsibility for what happens on their platforms. "(…) *how will you protect users' data? How do you proactively intend to stop harmful conduct instead of being forced to respond to it years or months later?"*[85] In rhetoric at least, Zuckerberg addressed the need for more human accountability over machines. "*We need to take a more proactive role and a broader view of our responsibility. It is not enough just to build tools, we have to make sure they are used for good*."[86]

3.2. How discourse was divided along party lines

While members from both parties raised the issue of public trust at the Facebook hearings, the rhetoric was clearly divided along partisan lines, with the ideological split more pronounced in the House of Representatives than in the Senate.[87] Some Republican members

[83] Idem. Sen. Grassley, C. p. 3 and p. 5
[84] Idem Sen. Nelson, B. p. 7
[85] Idem, p. 3
[86] Idem, Zuckerberg, M. p. 17
[87] The Senate with its longer terms of office has a tradition of more civil debate than the House of Representatives, See for example: http://www.gvptsites.umd.edu/uslaner/senate.pdf

tried to downplay the Trump campaign's ties to the Cambridge Analytica scandal, by equating it to the Obama campaign's successful use of social media microtargeting in 2008 and 2012. According to Senator Tom Tillis: "*Blame the Obama campaign for exploiting user data (…) take away the partisan rhetoric that is going on like this is only a Republican issue*."[88] There is an important distinction however: Obama's campaign did not use data that was inappropriately acquired. But Tillis may have a point that the Cambridge Analytica scandal was overblown. The ad consulting work that Facebook did with the Trump campaign in 2016, using legitimately acquired data, was arguably just as problematic from an ethical perspective, by heavily targeting potential supporters in swing states with divisive and anti-immigrant messages.[89] Equally troubling is Facebook unwittingly offering up its algorithms for money to Russian election interference efforts.

The ethical domain that caused the greatest partisan divide among legislators was the issue of justice, or whether Facebook users have equal opportunities on the platform, with accusations of algorithm bias split along party lines. A growing body of critical AI research has examined the question of algorithm bias, both in the statistical sense and the wider normative sense. Systemic weaknesses that can ingrain bias in the development of algorithms include 1) limited or incomplete test data, 2) faulty calculations and 3) a lack of diversity in the tech industry that limits the perspectives of the developers themselves.[90] Since machines are a reflection of the humans who create them, the question of whether they can ever be neutral decision-makers is therefore central to ethical discussions about their use.

---

[88] U.S. Senate, Sen. Tillis, T. p. 137

[89] Trump campaign manager Brad Parscale credits Facebook advertising for its winning strategy, and a Washington Post report found that Facebook consulting played a much bigger role in the Trump campaign's microtargeting efforts than Cambridge Analytica. In: Bump, P. (March 22, 2018).

[90] Campolo, A et al (2017) pp. 13-20, O'Neill, C (2016)

Republicans in the House of Representatives were particularly strident with accusations against Facebook of discriminating against evangelical Christian and conservative voices. In particular, several members denounced the platform for suspending the account of Diamond and Silk, two African-American sisters and fervent Trump supporters, for violating Facebook's community standards. Zuckerberg apologized for what he called a "mistake"[91], and the sisters' Facebook page, with its 1.2 million followers, was quickly restored. Interestingly however, the Fox News streaming service would later part ways with the two sisters after a series of false and misleading statements about the coronavirus that were apparently too controversial, even for the right-wing network.[92]

Zuckerberg acknowledged in the 2018 hearings that Silicon Valley "*is an extremely left leaning place (…) that is actually a concern that I have and that I try to root out in the company in making sure that we do not have bias in the work that we do*."[93] He went one step further by ordering an audit one month later to investigate claims of anti-Conservative bias from Facebook, led by a former Republican senator. The audit report, released in 2019, ended up displeasing everyone, angering Republicans for failing to clearly conclude that Facebook was biased against them, and angering Democrats for airing Conservative grievances without rebuttal and with no evidence to support them.[94] However, Zuckerberg did acquiesce to a Conservative demand that would later prove to be consequential: loosening restrictions on permissible content for political ads on Facebook. (More on this later in the paper). Facebook's efforts to placate Republicans were likely driven by profits rather than ethical concerns over equity, particularly as the GOP

[91] U.S. House of Representatives p. 57
[92] Cilizza, C. (April 20, 2020)
[93] Senate hearings, Zuckerberg, M. p. 49
[94] Overly, S. (Aug.20, 2019) and Diresta, R. (Aug. 21, 2019)

held the majority in both chambers of congress at the time he ordered the audit, and spent twice as much on election ads in 2016.[95] Furthermore, the accusation of anti-Conservative bias is difficult to sustain, given that 9 of the top 10 most active Facebook pages during the 2020 election campaign belonged to Conservative or pro-Trump voices.[96]

The charge of Facebook algorithm bias against minorities, levelled by Democrats, is more problematic from an empirical perspective. A 2016 ProPublica investigative report found that advertisers could deliberately exclude racial and ethnic groups from targeted advertising based on user profile categories developed by Facebook, a practice prohibited by federal law for housing and employment opportunities.[97] While the company settled a related civil suit and removed offensive exclusion categories from its advertising options, subsequent research found the company's new AI-based system that was meant to improve audience diversity for housing and employment ads had many of the same problems as its predecessor, with discrimination against certain users based on age and gender.[98] This reflects the risk with machine learning that decisions made on a discriminatory basis can be compounded, because algorithms make decisions based on historical patterns of practice.[99] (For example, Amazon discontinued an AI recruitment screening tool that discriminated against women as potential job candidates, based on historical data of Amazon's predominantly male workforce that was used to develop the algorithm.)[100] An independent civil rights audit commissioned by Facebook in response to public pressure was even more scathing. While it applauded the company for its efforts to increase the diversity of its workforce, it found that Facebook had not gone far enough to address

[95] Bump, P. (2020)
[96] Heilwell, R. (Sept. 2, 2020)
[97] Angwin, J. (Oct. 28. 2016)
[98] Kofman, A., Tobin, A. (2019)
[99] O'Neill, C. (2016)
[100] Dastin, J. (Oct. 10, 2018)

algorithmic bias, and failed to enforce its community standards, allowing the platform to facilitate the spread of hate against Muslims, Jews, and other targeted groups.[101] In particular, auditors took aim at Zuckerberg for taking a "*selective view of free expression*", by allowing rampant misinformation from politicians, including from the former President himself, to trample minority rights and facilitate voter suppression: "*Elevating free expression is a good thing, but it should apply to everyone. When it means that powerful politicians do not have to abide by the same rules that everyone else does, a hierarchy of speech is created that privileges certain voices over less powerful voices. The prioritization of free expression over all other values, such as equality and non-discrimination, is deeply troubling to the Auditors."*[102]

Finally, Democrats and Republicans in the House and Senate were divided over whether and how to regulate Facebook. Democrats argued that government regulation was long overdue, and that Zuckerberg's promises to "do better" were insufficient to ensure more ethical business decisions. "*Your business model is to monetize user information to maximize profit over privacy. And unless there are specific rules and requirements enforced by an outside agency, I have no assurance that these kinds of vague commitments are going to produce action*." [103] As proponents of small government and deregulation, Republican lawmakers leaned towards leaving Facebook alone or breaking up its quasi-monopoly, based on economic rather than ethical arguments. Some GOP legislators argued that regulation might only serve to further cement Facebook's market dominance, by adding new regulatory hurdles that would be too expensive for upstart competitors. Early reports on the impact of European privacy regulation

[101] Civil Rights Audit Final Report (2020) pp.6-8.
[102] Ibid, p. 8
[103] U.S. Senate, Blumenthal, R. p. 46

suggest there may be merits to this argument, as big tech giants such as Facebook and Google may have further consolidated their market positions in Europe as a result.[104]

3.3) What issues were missed by legislators

Non-maleficence: While legislators covered a lot of ground on the topic of non-maleficence, there was very little discussion on the physiological impact of Facebook's design on the mental health of individuals. Only two brief questions were raised about the impact of Facebook and social media on the well-being of teenagers, with only a passing reference to the platform's dopamine feedback loop design.[105]

There is now considerable research on what Citizen Lab Director Ron Diebert rightly calls "addiction machines".[106] A worldwide study asked students to go a full 24 hours without social media and they all used similar terminology to describe their reactions: "in withdrawal, frantically craving, extremely antsy, miserable and jittery."[107] This is not just hyperbole. A clinical research study, based on MRI brain scans, found that the constant stream of likes and shares from Facebook and other social media networks affect the brain's reward centre (the nucleus accumbens) by triggering the same kind of chemical reaction as other drugs, such as cocaine.[108] Another study found that frequent users of social media have greater activation in their amygdala, the part of the brain connected with impulsive behaviour and problematic substance use.[109] While not officially classified as a mental disorder, social media addiction nevertheless merits a lengthy entry for those seeking treatment from the U.S. Addiction

---

[104] Kostov, N. and Schechner, S. (June 17, 2019)
[105] Idem. Questions raised by Sasse, B. and Hassan, M. , pp.103 and 133
[106] Diebert, R. p. 29
[107] University of Maryland (2010), Diebert, R. p.30
[108] Dallas, M. (2013)
[109] Ghose, T. (2015)

Center[110], and the Canadian Mental Health Association cites several studies warning that patterns of continued use are "capable of altering the mood, motivation, concentration and producing a dissociating and disinhibiting experience for users."[111]

Zuckerberg claims active engagement on Facebook is in fact beneficent, and associated with positive mental health. "*When people are interacting with each other and posting and basically building relationships that is both correlated with higher measures of well-being, health, happiness, not feeling lonely, and that ends up being better for the business than when they are doing lower-value things like just passively consuming content*."[112] In other words, keeping people engaged on the platform for as long as possible is good for mental health AND company profits, according to Zuckerberg. However, his conclusions about positive mental health are based on internal Facebook research that is itself highly controversial.[113] In 2012, Facebook conducted an experiment on nearly 700,000 Facebook users to see if their mood could be altered by presenting different types of content in their news feeds. The study found that emotional states could be transferred without direct interaction between people, and that by manipulating the amount of positive or negative content on their news feeds, Facebook could "lead people to experience the same emotions without their awareness".[114] The so-called "emotional contagion" experiment, conducted without explicit user consent, may have been legal strictly speaking according to the company's terms of service, but was certainly unethical from an academic perspective, and evidence that Facebook was prepared to treat its users like lab rats.

---

[110] Addiction Centre: https://www.addictioncenter.com/drugs/social-media-addiction/
[111] Canadian Mental Health Association - https://ontario.cmha.ca/documents/addictions-and-problematic-internet-use/
[112] U.S. Senate, p. 134
[113] Meyer, R. (June 28, 2014)
[114] Kramer, A. et al (2014) p. 8788

Explicability: While the transparency of Facebook's user agreement was widely discussed, there was little probing from legislators on the transparency or explicability of the algorithms themselves. For example, how exactly do items appear on the individual newsfeeds of Facebook users? How are these items ranked and prioritized? And what effects are these decisions from machines having on us as individuals, and on broader society? (The emotional contagion experiment certainly proves that what appears in our news feed can have a physiological effect on us).

Democratic Senator Gary Peters was one of the few legislators to question the ranking decisions of Facebook's algorithms, asking why news feeds from friends were no longer being presented in chronological order, "*leaving people perplexed.(…) What else are your algorithms doing?"* he asked. Zuckerberg acknowledged that *"(r)ight now, a lot of our AI systems make decisions in ways that people do not really understand,*" and that AI transparency was going to be "*a very central question (…) over the next decade and beyond*."[115] However, there is a more immediate need to understand how the platform is engineering the social interactions of its users, and explanations to date about Facebook algorithmic logic are not always reassuring. In answer to a follow-up written question, Facebook replied that its newsfeed "*considers thousands of signals to surface the content that's most relevant to each person who uses Facebook*", ranking items for each individual based on the available inventory of stories, the computation of relevance scores, and predictions of what users are most likely to engage with.[116] Facebook also makes clear that humans are excluded from the decision-making process: *"Our employees don't determine the ranking of any specific piece of content*."[117] One can safely conclude however that

[115] U.S. Senate, Peters, G. and Zuckerberg, M. pp.121-122.
[116] U.S. Senate, p. 191.
[117] Idem

the average person does not understand how Facebook ranking algorithms work. In fact, many people do not even know that they exist. According to the Pew Research Centre, the majority of American Facebook users are unaware that machines are controlling what they see on their newsfeeds.[118]

In "Black Box Society", Frank Pasquale describes the contemporary world as a "one-way mirror", in which important corporate actors such as Facebook have unprecedented knowledge of the minutiae of our daily lives, while we know little to nothing about how they use this knowledge to influence the important decisions that we— and they— make.[119] Facebook algorithms determine what we get to see in our news feeds and when, leading to what the head of the Oxford Information Lab calls a "deeper, scarier, more insidious problem: we now exist in these curated environments, where we never see anything outside our own bubble (…) and we don't realize how curated they are."[120] Algorithm explicability is also central to the question of human autonomy in AI ethics: we cannot determine what decisions we want to delegate to machines if we do not consciously know what decisions machines are making, and on what basis these decisions are being made. The opacity of Facebook algorithms therefore leads to a loss of human control, a risk clearly articulated in the AI4People Framework.

Explicability of the polarization effect: While several members raised concerns about the societal impact of social media, there was surprisingly little discussion in the legislative hearings – save for a written question – about *how* Facebook algorithms enable polarization. A common critique of social media is that the most sensational or provocative material often tends

[118] Gramlich, J. (May 16, 2019)
[119] Pasquale, F. (2015) p. 5.
[120] Emily Taylor, chief executive of Oxford Information Lab, quoted in Hern, A. (2017)

to spread the fastest, due to algorithms that prioritize user engagement.[121] This polarizing effect is compounded by trapping users in echo chambers of like-minded people who are sealed off from other viewpoints, thereby increasing ideological divides.[122] While social media is not the only cause of political polarization, there is convincing empirical evidence of a strong association. The Pew Research Centre found that the divergence of opinions between Americans who identify as Democrat and Republican increased dramatically between 1994 and 2017, after social media entered the fray as a primary source of news for Americans.[123] Zuckerberg was asked in a written follow-up question by Senator Patrick Healey if his company was doing something about its algorithms to address the problem of polarization.[124]

Facebook replied that while it was "*keenly aware*" of concern that the platform was contributing to polarization, it took no direct responsibility, arguing that Facebook was simply a "*distribution platform that reflects the conversations, including polarized ones already taking place in society*."[125] The company further stated that data was mixed on what causes polarization and echo chambers, including its own research, which found that "*most people on Facebook have at least some friends who claim an opposing political ideology*". However, independent studies that have attempted to refute the echo chamber effect have not been particularly convincing. For example, a U.K. study found that people with a diverse media diet were able to avoid echo chambers online, but the study, based on a one-time survey, does not negate the fact that the effect still exists for a sizeable number of people.[126] Another empirical study that takes issue with Sunstein's theory suggests that the dynamics of online debates could be more aptly

[121] For example, Bucher, T. (2012) p. 1164
[122] Sunstein, C. (2011, 2017)
[123] Pew Research Centre (Oct. 20, 2017)
[124] Senate hearing transcripts, written question from Sen. Patrick Leahy, p. 322.
[125] Idem, Facebook reply, p. 322.
[126] Dubois, E., & Blank, G. (2018) p. 729

described by the logic of 'trench warfare', in which opinions are reinforced through contradiction as well as confirmation.[127] However, this merely reinforces the notion that social media amplifies "angry" opinions and can help to further polarize debate.

Facebook's defense of its algorithms raises interesting questions of explicability (in the sense of accountability), by highlighting contradictions between Zuckerberg's rhetoric of taking a "broader view" of his responsibility, and the company's statement that it is merely a distribution platform that is acting as a mirror of society. However, the algorithms driving content placement on Facebook are neither passive nor neutral. They actively determine what individuals get to see in their news feeds, and in the words of Facebook's former Chief Security Officer, Alex Stamos: "*Nobody of substance at the big companies thinks of algorithms as neutral. Nobody is not aware of the risks*."[128] Ironically, this statement, made prior to the Cambridge Analytica scandal, was meant to be an argument for the self-regulation of AI – as according to Stamos, big tech companies were the only ones with the technological knowledge to ensure ethical development of algorithms. However, according to the Wall Street Journal, Facebook executives ignored internal research in 2018 that showed the platform's algorithms were feeding users more and more divisive content to keep them online, potentially fueling extremism. "Our algorithms exploit the human brain's attraction to divisiveness," read a slide from the internal presentation, that warned of trouble ahead "if left unchecked."[129]

To summarize the hearings analysis, U.S. legislators were able to identify many ethical challenges with Facebook's business model aligned with four of the five principles of AI ethics (namely, non-maleficence, autonomy, justice and explicability). However, they were deeply

[127] Karlsen, R., et al. (2017), p. 257
[128] Tweet from Alex Stamos, (October 7, 2017)
[129] Horwitz, J. and Seetharamn, D. (May 26, 2020)

divided along party lines about how to interpret issues of justice such as algorithm bias, and on whether or how to regulate Facebook. In addition, legislators demonstrated only a superficial knowledge of how Facebook's ranking algorithms actually work, or the full extent of their impact on individuals and society – namely that they are opaque, they are designed to keep users "hooked", and the average user has no idea how intelligent machines are determining what they see on their news feeds. This amounts to a loss of human control, with a detrimental effect on societal cohesion, two further manifestations of unethical AI. The analysis also demonstrates the gulf between Zuckerberg's Pollyannaish rhetoric in the hearing chambers and his company's unethical policies and practices.

4. POLICY CONSIDERATIONS

During the two days of hearings, Zuckerberg repeatedly promised to address the two main "harms" – or breaches of the non-maleficence principle - that had brought him before the committee in the first place: Facebook's failure to protect user privacy (e.g. the Cambridge Analytica scandal), and the social network's failure to prevent the abuse of its platform by malicious actors (e.g. the Russian election interference scandal). This section examines what steps were taken by Facebook to address these harms, why they did not go far enough, and what lessons can be learned by policy-makers as a result. It also explores potential pathways for the regulation of social media and its algorithms to protect the public interest moving forward.

4.1. Steps taken by Facebook to improve privacy and transparency

At the 2018 hearings, Zuckerberg promised to better protect people's information on Facebook and to improve transparency for users.[130] However, the company's actions to date have either been reactive, self-serving or have not gone far enough. First, the company was forced to

[130] U.S. Senate, pp. 10-11

make changes in order to comply with the European Union's General Data Protection Regulation (GDPR), enacted in May 2018. Though drafted and passed in the EU, the GDPR imposes strict new privacy protection and transparency requirements for all companies collecting data from Europeans regardless of whether the organization is based in Europe or not.[131] As a result, Facebook made modifications to simplify its lengthy user consent forms, improve transparency features on targeted advertising, and add new privacy options for users across the board internationally, though not all of the GDPR-imposed changes apply in the U.S. or Canada. [132] (For example, the right to "erasure" which forces a company to delete anything that anyone has ever said about you does not exist in North America.)

Second, some of Facebook's actions on privacy were self-serving and primarily aimed at improving the company's bottom line. In order to "*make sure that what happened with Kogan and Cambridge Analytica does not happen again,"* Zuckerberg vowed to crack down on third party applications to prevent them from selling or storing the data of Facebook users, promising "*to take a much more proactive approach (going forward) and do regular spot checks*."[133] However, since data is the new "oil" of modern times, and user data is Facebook's single most valuable commodity, cutting off the supply to third parties may be more about increasing Facebook profits than about ethical concern for user privacy. After all, Cambridge Analytica was accused of performing essentially the same service with user data – i.e. creating granular user profiles for the micro-targeting of campaign propaganda – that Facebook offers at a premium for political parties. Why outsource to others, what you could do yourself for bigger revenues?[134]

---

[131] Information on GDPR at: https://gdpr.eu/what-is-gdpr/
[132] Recap of privacy improvements in: Facebook (Jan. 28, 2021)
[133] U.S. Senate, Zuckerberg, M. pp.10 and 14
[134] The issue of Facebook's "embedded" campaign marketing consultants was discussed in House of Representatives hearings, p.54.

Finally, Facebook's quest to improve transparency does not go far enough. Over the course of the 2018 congressional hearings, Zuckerberg repeatedly argued that the company's aim was to give users full control over their own data, stressing that "*users get to choose what they post on Facebook and who gets to sees it – and are explicitly asked for permission every time they share information on Facebook."*[135] However, Zuckerberg was evasive throughout the hearings about the full extent of Facebook's surveillance apparatus. Facebook does not just glean information from what you proactively share, but collects your personal data based on where you go, what devices you use, and how you surf online – even when logged out of the Facebook platform. In fact, this is how Facebook algorithms most accurately determine who we are, what we like and what triggers our attention – not by what we *say* but by how we *behave* online.[136] When asked in the House of Representatives if he would commit to changing all of Facebook's default user settings to minimize the collection of user data to the greatest extent possible, Zuckerberg refused to answer.[137]

It should therefore come as no great surprise that the issue to most rankle the usually unflappable Zuckerberg was Apple's recent decision to change default options on its iPhones and other devices. The changes would require users to "opt in" – or physically change the default settings on their new devices – in order to allow companies such as Facebook to track them across devices and applications. As Apple CEO Tim Cook explained, in an apparent jab at Zuckerberg: "Technology does not need vast troves of personal data, stitched together across dozens of websites and apps, in order to succeed." [138] The move has triggered an online battle

[135] U.S. Senate, p. 11.
[136] Howard, P. (2020) pp. 3 and 11.
[137] U.S. House of Representatives, exchange with Pallone, F. p.19
[138] Edelman, G. (2021, Jan. 28)

between the two tech giants, with Facebook reportedly considering a lawsuit against Apple on the grounds of anticompetitive practices.[139] However, it also signals the value Facebook places on the ability to track and surveil users without their explicit and proactive consent beyond the defined parameters of the platform.

4.2 Steps taken by Facebook to control the spread of misinformation

Under the harsh political and media spotlight of the 2018 legislative hearings, Zuckerberg announced a series of measures to help fight rampant disinformation and extremism on the social media platform. This included hiring an army of content moderators and third-party fact checkers that grew from 15,000 in early 2018 to more than 30,000 in preparation for the U.S. election campaign of 2020.[140] The company also set out to fight fire with fire, by using artificial intelligence as the first line of attack against malicious digital actors, a "beneficent" use of technology from an ethical perspective. Facebook claims it has made successful use of advanced AI tools to proactively police the platform, including identifying and halting ISIS and Al-Qaeda terrorism recruitment efforts, and removing thousands of fake accounts and web pages linked to Russia's Internet Research Agency.[141]

The technology has the benefit of being able to scan data from more than two billion users and identify red flags faster than any number of humans possibly could. However, AI tools are not a panacea, and challenges to date have been two-fold. First, bad actors such as Russia are constantly improving their means of deception, and remain active competitors in the digital "arms' race". Zuckerberg admitted in the 2018 hearings that "*as long as there are people sitting*

[139] Isaac, M., & Wakabayashi, D. (2021, January 28)
[140] U.S. Senate, p. 11 and https://about.fb.com/actions/promoting-safety-and-expression/
[141] Idem pp.10-11

in Russia whose job it is to try to interfere with elections around the world, this is going to be an ongoing conflict."[142] Second, hate is not always easy for non-sentient beings to detect. The line between legitimate political discourse and hate speech can be difficult to identify, especially when relying on machines to make the initial discovery. Zuckerberg warned legislators that the technology was still 5 to 10 years away from accurately deciphering the linguistic nuances of hate speech.[143] This has led to flawed decisions from Facebook algorithms, such as flagging posts that are condemning hate and violence while allowing racist dog whistles and more subtle incitements of hatred to stand and spread online.

However, the most problematic shortcomings of Facebook to stop the spread of disinformation and extremism may be due to company policy choices, rather than machine deficiencies. Take the issue of political advertising on Facebook for example. To make amends for abetting Russian election interference, Zuckerberg agreed to label all political and issues-based ads on the platform as well as verify the identity and location of those buying the ads, in advance of proposed new U.S. Senate legislation, the "Honest Ads Act".[144] However, this does not prevent foreign operatives from using a domestic buyer to purchase political ads, and it certainly does not prevent the spreading of political lies. In fact, later in 2018, Facebook modified its advertising policies to exempt political figures from the company's rules against ads that contain false or misleading information.[145] This led to a viral exchange in 2019 congressional hearings between Zuckerberg and Representative Alexandria Orcasio-Cortez, who asked if she could run advertisements targeted at Republicans in the primaries saying they had

---

[142] Idem. p. 113.
[143] Idem, p.18, Zuckerberg insisted he "*was not waiting for legislation to act*" on changes to political advertising.
[144] Idem, p. 13
[145] Durkee, A. (Oct. 2019)

voted for the Green New Deal.[146] Senator Elizabeth Warren took the argument a step further, by having her campaign create its own false Facebook ad claiming that Zuckerberg endorsed Trump. "*They've decided to let political figures lie to you – even about Facebook itself – while their executives and their investors get even richer off the ads containing these lies,*" Warren wrote on Twitter.[147]

Zuckerberg defended his company's ad policy by arguing that "*in a democracy, I believe that people should be able to see for themselves what politicians (…) are saying and judge their character for themselves."*[148] However, Facebook's "free pass" for politicians extends well beyond advertisements. In late 2018, Facebook created a "political whitelist" of about 112,000 accounts belonging to government officials and candidates. Those on the list could post freely without being fact checked or facing penalties, such as reduced algorithmic reach, for the spread of misinformation. Internal Facebook research found that people were more likely to believe false information if it was shared by a politician from the party they voted for. In 2019, the researchers in question made an internal pitch to dissolve the list, arguing it was both maleficent and inequitable: "*Allowing whitelisted content to evade our typical misinformation interventions privileges the very content that has an increased likelihood of deceiving people as to its trustworthiness and utility*," their presentation summarized.[149] Their proposal was rejected by Facebook executives.

Organic sharing remains one of the most important means of spreading misinformation on Facebook. The bulk of Russia's disinformation efforts were not conducted through

---

[146] Congressional hearing (Oct. 23, 2019) at 2:56./www.youtube.com/watch?v=G272R50v6ww&t=251s
[147] Durkee, A. (Oct., 2019).
[148] See 126, (reply to Orcasio-Cortez) at 4:00.
[149] Heath, A. (Nov. 24, 2020)

advertising but by actively participating on social media platforms to spread fake news and amplify division across enemy lines.[150] The end game is not just to elect candidates who are more favourable to the Kremlin but to destabilize adversaries by undermining confidence in democratic institutions. It's why Russia's cyber campaigns in the U.S. did not subside after the election of Donald Trump, but infiltrated online chats to exploit societal cleavages and amplify divisive messages over issues such as race, immigration and religion[151]. Since 2016, Facebook has vastly improved its capacity to identify and remove Russian cyber-propaganda, but by 2020, Russian help was no longer required to enflame tensions and manipulate voters ahead of the U.S. election. The Trump campaign and its supporters were fully capable of manufacturing, distributing and targeting lies by themselves, with Facebook's tacit permission. A U.S. study found the level of engagement on Facebook with verifiably false content produced domestically in 2020 had increased by 102% since the run-up to the 2016 election, while Facebook users engaged three times more frequently with pro-Trump "news" sites that manipulate the truth.[152] In short, the lie machines perfected by authoritarians in the Kremlin were emulated by the political mainstream of the most powerful democracy on earth.[153]

In the summer of 2020, Facebook promised new measures to limit the spread of "hateful language" and to improve labeling of political posts after advertisers threatened to boycott the platform. [154] The changes meant that Donald Trump and other important political figures could continue to post divisive disinformation in violation of the platform's community standards, but with a label that the content was "newsworthy" enough to remain. By this point however, with or

[150] Bump, P. (2018), and U.S. Justice Dept. (2018)
[151] Ample empirical evidence is provided in Barrett, P. et al (2018)
[152] Kornbluh, K. (2020). Note that left-wing sites in the U.S. were also found to peddle misinformation but on a much smaller scale.
[153] Lie machines in authoritarian and democratic countries documented in Howard, P. (2020)
[154] Isaac, M. Frenkel, S. (June 26, 2020)

without a disclaimer, large swaths of the general population in America and elsewhere were ready and willing to embrace lies and conspiracy theories, indoctrinated by years of algorithmic overexposure.

Social media is not the only reason for the rise of post-truthism, polarization and extremism. In the U.S., the decline of the middle class, barriers to higher education, and a conservative cable and talk radio information ecosystem have all played an enabling role.[155] But social media has had a steroid effect on the distribution of lies, and it does this by hijacking our cognitive biases. In the age of information overload, we tend to trust friends and family as good sources of information, and rely on heuristics for evaluating the truth. According to Oxford researchers, the most important way that misinformation is spread is from average citizens who make the mistake of sharing fake news on their personal feeds.[156] Once a lie is accepted, cognitive dissonance theory holds that people want to feel they are right at all costs, even when confronted with contradictory information. Irrational views are further reinforced when a large number of other people share them;[157] on Facebook, such views can be magnified exponentially by bots, trolls and algorithms seeking to grab the attention of like-minded people.

In the midst of a global pandemic, disinformation has come at a deadly cost. The spread of conspiracy theories on Facebook and other social media platforms has led to a mistrust of science that has moved beyond the extreme fringes of society, and into the mainstream. A Léger poll last October found that 1 in 5 Quebeckers believed in falsehoods or conspiracies about the coronavirus[158], while according to a Civiqs poll in the U.S., a majority of Republican supporters

[155] Luce, E. (2017)
[156] Howard, P. (2020), p.1
[157] McIntyre, L. (2017) summarizing the work of Leon Festinger and other behavioural scientists.
[158] Péloquin, T. (Oct. 24, 2020)

believe at least part of the QAnon theory[159]. A British study offers empirical evidence of Facebook's role in enabling QAnon's spread of false claims about the severity of COVID-19, anti-mask propaganda, and baseless theories on the origin of the virus. The analysis of over 200,000 Facebook posts found engagement with QAnon almost doubled between April and August 2020, with influential "superspreaders" sharing introductory materials, posting videos, and instilling a "do-it-yourself ethos" to lead followers down the indoctrination rabbit hole.[160]

Assault on the capital: the failure of self-regulation and Frankenstein fully unleashed

During the 2020 U.S. election campaign, Zuckerberg tried to play both sides, by giving carte blanche to lying politicians on the one hand (a nod to the President and his supporters), and supporting civic engagement on the other, by adding education tools on Facebook and donating $400 million to "get-out-the-vote" grassroots organizations (a nod to the Democrats). In the tumultuous post-election period, when Trump refused to concede the election, Facebook and Twitter added warning labels on incendiary false claims of election fraud from Trump and his political supporters. The most dramatic action was only taken however after January 6, when a mob egged on by the former President, stormed the Capitol building during the vote to certify Joe Biden's election, causing 5 deaths. The following day, Facebook suspended Donald Trump's account, with Mark Zuckerberg posting that "*the risks of allowing the President to continue using our service during this period are simply too great*".[161] Twitter quickly followed suit by shutting down @realDonaldTrump, the president's favourite online megaphone, and banning all of his associated accounts.

---

[159] Civiqs (Sept. 2, 2020)
[160] O'Connor et al (2020)
[161] See post at https://about.fb.com/news/2021/01/responding-to-the-violence-in-washington-dc/

Other big tech companies joined the fray, with Google and Apple both removing Parler and other extremist social media alternatives from their application offerings until the platforms could be moderated.[162] In the end, big business took action to limit ideas that were simply too toxic for mainstream consumers. The free market eventually worked to correct the negative externalities it helped to create, only much too late. It was in the words of one commentator like "shutting the barn door long after the horse had already bolted."[163] In fact, multiple reports detailed the role that the Facebook platform played in facilitating the January 6 "Stop the Steal" rally in Washington and connecting extremists online in preparation for the big event.[164]

Despite Facebook's most recent efforts, it has become a highly effective platform for spreading hate. While other platforms such as 4Chan (which later became 8kun) are more blatant niche promoters of violent white extremism, Facebook has two things going for it that 8kun does not: its unparalleled audience reach and algorithms that personalize the user experience. This means that no two people have exactly the same news feed, making the spread of hate less openly obvious and more insidious than on digital billboard sites such as 8kun. As Jonathan Geltzer, a former white house counterterrorism expert points out, Facebook "is not a filter bubble; it's a filter shroud."[165] Mary McCord, the legal director at the Institute for Constitutional Advocacy and Protection at Georgetown Law explains the platform's network effects: "In every situation of extremist violence we've looked into, we've found Facebook postings. And that reaches tons of people. The broad reach is what brings people into the fold and normalizes extremism and makes it mainstream."[166]

---

[162] Goldberg, M. (Jan 11, 2021)
[163] Robert Reich on Twitter (Jan. 7, 2021)
[164] Examples include Dwoskin, E. (Jan. 13, 2021) and Deryesh, I. (Jan. 16, 2021)
[165] Quoted in Lafrance, A. (2020)
[166] Idem

4.3. Moving forward: pathways for ethical regulation of social media

Efforts to regulate Facebook to date have been largely insufficient. In 2019, Facebook received a $5 billion fine for the Cambridge Analytica privacy breach – the largest fine in the F.T.C.'s history.[167] However, this represents less than the company earns in a month, and did not make a dent in the company's growing profit margin.[168] The F.T.C's recent bid to break up Facebook for its anti-competitive practices, hailed as long overdue by some critics, may diminish the oversized market weight of the company. However, a market-based approach to regulation does not address the larger ethical problems with social media and its ill-effects on individuals and society. For example, new competitors are more than happy to pick up where Facebook left off in the peddling of lies and conspiracies; Substack is just the latest platform hoping to become the next virtual megaphone for Trump.[169]

The problem with social media is fundamentally one of accountability (i.e. the explicability principle) – and the need for humans to be responsible for the technology they manage and the algorithms they produce. Facebook is not just a tech company but the "world's biggest publisher" – with about 140 million Americans getting their news predominantly from the platform.[170] However, the company does not face the same legal responsibilities as newspapers or broadcasters for the content it distributes, as Section 230 of the U.S. Communications Decency Act exempts internet companies from legal liability.[171] The internet has thus eliminated the traditional "gatekeeper" role of the media for widely distributed information, and along with it, responsibility for information quality.

---

[167] Federal Trade Commission (July 24, 2019)
[168] See in Statista. https://www.statista.com/statistics/268604/annual-revenue-of-facebook/
[169] Swisher, K. (Jan. 25, 2020)
[170] Senate hearings, citing Pew Research.
[171] Under the USMCA, similar rules apply in Canada.

Just as airwaves are considered a public good, the internet needs to be treated as a public good, with corresponding standards and rules to protect against market failures and abuse. Big tech needs to fix what it helped to destroy, by restoring the business model of an accountable press. This means not just better labeling and promotion of real news through algorithm adjustments, but contributing to its funding as well. To this end, a levy could be collected from social media companies to fund responsible journalism, and perhaps even civic education, to encourage broader critical thinking in citizens. Australia has taken a bold first step, by introducing legislation to force Facebook and Google to pay broadcasters and content producers for their content.[172] (Facebook initially punished Australia last month by temporarily blocking Australian news organizations on its platform but has since struck a deal with Canberra).

Free speech is a cornerstone of liberal democracy, and while the right to free speech – be it the First Amendment in the U.S. or the Charter of Rights and Freedoms in Canada – complicates the job of policing misinformation, it is possible to protect freedom of expression while demolishing the mechanisms for producing, distributing and marketing hate. German law requires social media companies to remove hate speech within 24 hours of it appearing online,[173] but in the U.S. such a law would be unconstitutional and in Canada it is considered impalatable by some experts[174]. Facebook's community standards already exceed legal requirements by preventing the posting and sharing of content such as the support of violent or criminal acts, harassment, false reports, and hate speech.[175] The problem has been two-fold: a) a lack of timely and consistent enforcement and b) an unequitable application of these standards for political

[172] Packham, C. (Feb. 11, 2021)
[173] Krunke, C. (2020) p. 154
[174] Curry, B. (Jan. 27, 2021)
[175] See at https://www.facebook.com/communitystandards/

expediency. A further challenge is the idea that private companies alone get to determine what information is acceptable, particularly in light of the central function that social media platforms now play in our social lives. From this perspective, a rebalancing of the relationship between ordinary citizens and the tech companies that collect data from them offers a promising alternative. Yale Professor Jack Balkin proposes the concept of "information fiduciaries", whereby big tech companies, in light of their special powers over average citizens, have legal obligations, just as doctors and lawyers do, toward their clients and the people they serve. This approach allows social media companies to face greater regulation in conformity with the First Amendment, while providing greater transparency and accountability for the public.[176]

Good regulation also requires strong institutions. In the age of the internet of things, in which ever increasing amounts of behavioural data can be collected from our watches, accessories and appliances, strong oversight will be imperative to protect people's right to privacy, give people control over their own data, and truly respect the principle of human autonomy. In Canada, a federally funded panel is recommending the creation of a new government regulator to ensure social media companies act responsibly, with the power to impose penalties such as fines and possibly jail time for breaches. [177] It says companies should be required to provide strong and transparent content moderation policies, explain their algorithms, and offer clear mechanisms for contesting their decisions, and follows the lead of a similar initiative in the U.K.[178] In the U.S., with the Democrats now in charge of both houses, there is a strong appetite to move forward with regulation after years of congressional stalling. This includes the creation of an independent Data Protection Agency, requiring tech companies to be

[176] Balkin, J. (2016) pp. 1186-87
[177] Curry, B. (Jan. 26, 2021)
[178] Ofcom (Dec. 15, 2020)

more accountable at the federal level for how they collect and use personal data, including providing more transparency for algorithms.[179]

The fundamental issue for some observers, is Facebook's algorithm-based business model. For Cass Sunstein, the quest to make content "relevant" for users is what is keeping us in echo chambers, and tearing society apart. Shoshanna Zuboff goes further, arguing that surveillance capitalism needs to be completely dismantled if democracy is to survive.[180] While social media companies are not likely to disappear overnight, with the right pressure from the public and regulators, they can certainly be detoxified. This includes modifications to algorithms to de-amplify the spread of sensationalist content, offering users alternative perspectives from validated sources to dissolve filter bubbles, and limiting the over-reach of Facebook and other tech giants, by making users opt in to share their data across devices and platforms.

CONCLUSION: LESSONS FOR THE AGE OF DISRUPTION

From 2004 to 2014, Facebook's motto has been "*to break things and move on*". It has more than lived up to this promise. Facebook and other social media giants have not just radically disrupted the publication and broadcast industries, they have re-engineered our social and political lives. The company that pledged to help "*build community and bring the world closer together*" allowed its algorithms to hack people's emotions, target the vulnerable, and become super-spreaders of lies, conspiracy theories, and anti-democratic propaganda – all while increasing its revenues by $30 billion over the past two years alone. As one Democratic House member put it, Facebook is "*always at the scene of the crime*".[181]

[179] Electronic Privacy Information Center (April 2020)
[180] Zuboff, S. (Jan. 21, 2021) – similar views are also shared by Diebert, R., and Howard, P.
[181] Rep. Greg Meeks quoted in Durkee, A. (Oct. 24, 2019)

The trouble with Facebook is a failure to self-regulate, but also a failure to anticipate. As Floridi et al point out, AI is not just another utility, but a powerful form of smart agency that must be steered from the onset to benefit society.[182] But as the Facebook experiment clearly demonstrates, industry cannot be trusted to ensure the ethical development of algorithms, and a soft law approach is no longer acceptable. Despite a corporate ethics code and improvements made to the platform since the Cambridge Analytica scandal, Facebook has failed to comply with four of the five key principles of ethical AI – namely non-maleficence, justice, autonomy and explicability – diluting the benefits it provides to society in the process (the fifth beneficence principal).

Legislators in the U.S. also deserve a share of the blame. While they have raised ethical concerns about Facebook in their political rhetoric, legislative action has been hindered by a lack of technical knowledge and mired in partisan politics. A new government in Washington offers new hope for corrective action, but the regulation of social media algorithms cannot be left to the whims of whatever party happens to be in power. Independent oversight is required to ensure ethical collection and use of data, and to set the boundaries for responsible algorithm development. As recent events at the Capitol have demonstrated, the stakes of unregulated technology are simply too high.

In his modernization theory, Ulrich Beck argues that we have made an epochal shift from the industrial age to the risk society era, in which scientists, technologists and corporations are treating the world as a laboratory. The innovations that are supposed to be improving our lives are dramatically altering our social and physical world, with many "out of control processes" and

[182] Floridi, L. et al, p. 698

"systemic unintended side effects" placing us at increased risk on a global scale.[183] The precautionary principle has gained traction internationally to mitigate public health and environmental threats from technology, even when scientific evidence about these threats is uncertain. The time has come to consider a precautionary approach for artificial intelligence in the information era. Two decades of unrestricted growth by giant tech companies, with Facebook leading among them, have damaged human brains and put democracy at risk.

The story of Facebook serves as a cautionary tale for what lies ahead. Social media algorithms are merely one of many applications of artificial intelligence, and as machines get smarter with less human control, futurists are predicting that the AI revolution will not be a single watershed event but a "cascade of ever bigger disruptions."[184] How prepared are policy-makers to anticipate these changes and mitigate their negative effects? And is a global solution required given that technology knows no borders?[185] These questions become all the more relevant as science fiction fantasies about technology are fast becoming reality: next-generation wearable devices currently under development can read your thoughts to command objects connected by the internet-of-things[186], and scientists believe the development of a "super-intelligent" machine, that is resistant to human control, is only a brief matter of time.[187]

,

[183] Beck, U. (1992)
[184] Harari, Y. (2018), p. 33
[185] Erdélyi.O.E. and Goldsmith, J.(2018) p.2
[186] Golembiewski, L. (2019)
[187] Bostrom, N. (2014) in his book on the topic

Bibliography

Angwin, J., Parris, T. (Oct. 28, 2016) Facebook Lets Advertisers Exclude Users by Race. *ProPublica*. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

Asimolar AI Principles (2017) Principles developed in conjunction with the Asimolar conference on Benevolent AI. *The Future of Life Institute*. https://futureoflife.org/ai-principles/

Balkin, J. (2016). Information fiduciaries and the First Amendment. *U.C. Davis Law Review*, *49*(4), 1183–. https://lawreview.law.ucdavis.edu/issues/49/4/Lecture/49-4_Balkin.pdf

Barrett, P., Wadhwa, T. and Baumann-Pauly,D. (2018) *Combating Russian Disinformation: The Case for Stepping Up the Fight Online*. NYU Stern Centre for Business and Human Rights, July 2018) https://issuu.com/nyusterncenterforbusinessandhumanri/docs/nyu_stern_cbhr_combating_russian_di/1.

Beck, U. (1992). *Risk Society: Towards a New Modernity*. Translated by Ritter, Mark. London: Sage Publications

Berinsky, A. J. (2017) *Rumours and Health Care Reform: Experiments in Political Misinformation*. British Journal of Political Science, 47

Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. http://ebookcentral.proquest.com/lib/ottawa/detail.action?docID=1743975.

Bradshaw, H. (2020). Sourcing and Automation of Political News and Information over Social Media in the United States, 2016-2018. *Political Communication*, *37*(2), 173–193. https://doi.org/10.1080/10584609.2019.1663322

Bücher, T (2012) Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society* 14(7): 1164–1180.

Bump, P. (March 22, 2018). All the Ways Trump's Campaign was Aided by Facebook Ranked by Importance. *Washington Post*. https://www.washingtonpost.com/news/politics/wp/2018/03/22/all-the-ways-trumps-campaign-was-aided-by-facebook-ranked-by-importance/

Campolo, A., Sanfilippo, M., Whittaker, M., Crawford, C. *AI Now 2017 Report*. AI Now Institute, New York University. https://ainowinstitute.org/AI_Now_2017_Report.pdf

Cilizza, C. (April 28, 2020) Fox's firing of Diamond & Silk isn't the problem. Fox's hiring of them is. *CNN*. *https://www.cnn.com/2020/04/28/politics/diamond-silk-fox-news-donald-trump/index.html*

Civiqs (Sept. 2, 2020) *Report : Americans Pessimistic on Time Frame for Coronavirus Recovery.* https://civiqs.com/reports/2020/9/2/report-americans-pessimistic-on-time-frame-for-coronavirus-recovery

Curry, B. (Jan. 27, 2021) Report Calls for Powerful New Federal Body to Regulate Social Media. *The Globe and Mail*. https://www.theglobeandmail.com/politics/article-report-calls-for-powerful-new-federal-body-to-regulate-social-media/

Dallas, M. (2013) Facebook Use Linked to Brains Reward System. *Retrieved from:* http://www.medicinenet.com/script/main/art.asp?articlekey=173165

Dastin, J. (Oct. 10, 2018) Amazon scraps secret AI recruiting tool that shows bias against women. *Reuters*. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Deibert, R. (2019). The Road to Digital Unfreedom: Three Painful Truths About Social Media. *Journal of Democracy*, (1), 25-39.

Deryesh, I. (Jan. 16, 2021) Despite Parler backlash, Facebook played huge role in fueling Capitol riot, watchdogs say. *Salon. https://www.salon.com/2021/01/16/despite-parler-backlash-facebook-played-huge-role-in-fueling-capitol-riot-watchdogs-say/*

Diresta, R. (Aug. 19, 2019) Facebook's anti-conservative bias audit results are here; its results are entirely unsurprising. *Slate*. https://slate.com/technology/2019/08/facebook-anti-conservative-bias-audit-results-kyl.html

Dubber, M., Pasquale, F., & Das, S. (2020). *The Oxford Handbook of Ehics of AI*. Oxford University Press;

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5), 729–745. https://doi.org/10.1080/1369118x.2018.142865

Dwoskin, E. (Jan. 13, 2021) Facebook's Sandberg deflected blame for Capitol riot, but new evidence shows how platform played role. *Washington Post*. https://www.washingtonpost.com/technology/2021/01/13/facebook-role-in-capitol-protest/

Edelman, G. (Jan. 28, 2021) Facebook and Apple Are Beefing Over the Future of the Internet, *Wired*. https://www.wired.com/story/facebook-apple-feud-over-privacy-internet-future/

Electronic Privacy Information Center (April 2020) *Grading on a Curve: Privacy Legislation in the 116th Congress*. https://epic.org/GradingOnACurve/EPIC-GradingOnACurve-Apr2020.pdf

Erdélyi.O.E. and Goldsmith, J. (2018). *Regulating Artificial Intelligence Proposal for a Global Solution*. In 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES'18), February 2–3, 2018, New Orleans, LA, https://doi.org/10.1145/3278721.3278731

Facebook (June, 2020) *Facebook's Civil Rights Audit – Final Report*. https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf

Facebook (Jan. 28, 2021). Recapping our privacy controls on data privacy day. https://about.fb.com/news/2021/01/recapping-our-privacy-controls-on-data-privacy-day/

Federal Trade Commission (July 24, 2019) *Federal Trade Commission Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook*. https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions

Federal Trade Commission (Dec. 9, 2020) *FTC Sues Facebook for Illegal Monopolization*. https://www.ftc.gov/news-events/press-releases/2020/12/ftc-sues-facebook-illegal-monopolization

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* (Dordrecht), 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Frischmann, B., & Selinger, E. (2018). *Re-Engineering Humanity*, Cambridge University Press.

Ghose, T. (2015). What Facebook Addiction Looks like in the Brain. *Life Science.* http://www.livescience.com/49585-facebook-addiction-viewed-brain.html

Goldberg, M. (2021) The Scary Power of the Companies That Finally Shut Trump Up. *The New York Times*, Jan. 11, 2021. https://www.nytimes.com/2021/01/11/opinion/twitter-facebook-trump-ban.html

Golembiewski, L. (2019) How Wearable AI Will Amplify Human Intelligence, *Harvard Business Review*, April 2019. https://hbr.org/2019/04/how-wearable-ai-will-amplify-human-intelligence

Grieco, E. (April 20, 2020) US newspapers have shed half of their employees since 2008, *Pew Research Centre*. https://www.pewresearch.org/fact-tank/2020/04/20/u-s-newsroom-employment-has-dropped-by-a-quarter-since-2008/

Harari, Y. (2018). *21 Lessons for the 21st Century*. Jonathan Cape.

Hagendorff, T. (2020). *The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines* (Dordrecht), 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Harrington, L. (1996). Ethics and Public Policy Analysis: Stakeholders' Interests and Regulatory Policy. Journal of Business Ethics, 15(4), 373-382. Retrieved December 24, 2020, from http://www.jstor.org/stable/25072761

Heath, A. (Nov. 24, 2020) Facebook Researchers Found Its 'Political Whitelist' Influenced Misinformation Spread. *The Information*. https://www.theinformation.com/articles/facebook-researchers-found-companys-political-whitelist-influenced-misinformation-spread

Heilweil, R. (Sept. 20, 2020) Right-wing media thrives on Facebook. Whether it rules is more complicated. *Vox Media. https://www.vox.com/recode/21419328/facebook-conservative-bias-right-wing-crowdtangle-election*

Hern, A. H (May 22, 2017) How social media filter bubbles and algorithms influence the election. *The Guardian.* http://www.theguardian.com/technology/2017/may/22/social-media-election-facebook-filter-bubbles

Horwitz, J. and Seetharamn, D. (May 26, 2020) Facebook Executives Shut Down Efforts to Make the Site Less Divisive. *Wall Street Journal.* https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499

House of Commons Digital, Culture, Media and Sport Committee. (February 2019) *Disinformation and 'fake news': Final Report.* Eighth Report of Session 2017–19, House of Commons, U.K. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

Howard, P. N. (2020). *Lie machines: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives*. New Haven: Yale University Press.

Isaac, M. and Frenkel, S. (June 26, 2020) Facebook Adds Labels for Some Posts as Advertisers Pull Back. *New York Times.* https://www.nytimes.com/2020/06/26/technology/facebook-labels-advertisers.html

Gramlich, J. (May 16, 2019) 10 Facts About America and Facebook. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2019/05/16/facts-about-americans-and-facebook/

Isaac, M. & Wakabayashi, D. (2021, January 28) Facebook Said to Consider Suing Apple Over App Store Practices. *New York Times*. https://www.nytimes.com/2021/01/28/technology/facebook-apple-app-store.html

Jobin, A., Ienca, M. & Vayena, E. The Global Landscape of AI Ethics Guidelines. *Nat Mach Intell* 1, 389–399 (2019). https://doi.org/10.1038/s42256-019-0088-2

Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication (London)*, *32*(3), 257–273. https://doi.org/10.1177/0267323117695734

Kofman, A. and Tobin, A. (Dec. 13, 2019) Facebook Ads Can Still Discriminate Against Women and Older Workers Despite a Civil Rights Settlement. *ProPublica.* https://www.propublica.org/article/facebook-ads-can-still-discriminate-against-women-and-older-workers-despite-a-civil-rights-settlement

Kornbluh, K. (2020) New Study by Digital New Deal Finds Engagement with Deceptive Outlets Higher on Facebook Today Than Run-up to 2016 Election. *German Marshall Plan Fund of the United States. https://www.gmfus.org/blog/2020/10/12/new-study-digital-new-deal-finds-engagement-deceptive-outlets-higher-facebook-today*

Kostov, N. and Schechner, S. (June 17, 2019) GDPR Has Been a Boon for Facebook and Google. *The Wall Street Journal. https://www.wsj.com/articles/gdpr-has-been-a-boon-for-google-and-facebook-11560789219*

Kramer, A., Guillorey, J., Hancock, J. (2014) Emotional Contagion Through Social Networks. *Proceedings of the National Academy of Sciences* Jun 2014, 111 (24) 8788-8790; DOI: 10.1073/pnas.1320040111

Krunke, C. "Artificial Intelligence in Social Media". In: Wischmeyer, T., & Rademacher, T. (2020). *Regulating Artificial Intelligence*. Springer International Publishing. Pp. 145-171. https://books-scholarsportal-info.proxy.bib.uottawa.ca/en/read?id=/ebooks/ebooks5/springer5/2019-12-31/1/9783030323615#page=1

Lafrance, A. (December, 2020) Facebook is a Doomsday Machine. *The Atlantic*. December 2020. https://www.theatlantic.com/technology/archive/2020/12/facebook-doomsday-machine/617384/

Luce, E. (2018) *The Retreat of Western Liberalism*. Grove Press.

Meyer, R. (June, 28, 2014) Everything We Know about Facebook's Emotional Contagion Experiment. *The Atlantic*. https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

McIntyre, L (2017) *Post-Truth*, MIT Press.

Mintzer ,A. (March 31, 2020) Paying Attention: the Attention Economy. *The Berkeley Economic Review* https://econreview.berkeley.edu/paying-attention-the-attention-economy/

Montreal Declaration for a Responsible Development of Artificial Intelligence (2018) *Université de Montréal*. https://www.montrealdeclaration-responsibleai.com/the-declaration

O'Connor, C. et al, (2020) *The Boom Before the Ban: QAnon and Facebook*. The Institute for Strategic Dialogue. https://www.isdglobal.org/isd-publications/the-boom-before-the-ban-qanon-and-facebook/

Ofcom (Dec.15. 2020) Ofcom to regulate harmful content online. https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/ofcom-to-regulate-harmful-content-online

O'Neill, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.

Office of the Privacy Commissioner of Canada (Aug. 2009) *Facebook Agrees to Address Privacy Commissioner's Concerns.* https://www.priv.gc.ca/en/opc-news/news-and-announcements/2009/nr-c_090827/

Organisation for Economic Cooperation and Development (June 11, 2010) *The Future of News in the Internet.* https://www.oecd.org/sti/ieconomy/oecdexaminesthefutureofnewsandtheinternet.htm

Overly, S. (Aug. 20, 2019) Conservatives, liberals, slam Facebook bias audit. *Politico*. https://www.politico.com/story/2019/08/20/facebook-audit-conservative-bias-amid-criticism-1469287

Packham, C. (Feb. 11, 2021) Australia to introduce landmark Google, Facebook legislation to parliament next week. *Reuters*. https://www.reuters.com/article/us-australia-media-google/australia-to-introduce-landmark-google-facebook-legislation-to-parliament-next-week.html.

Pariser, E (2011) *Filter Bubble. What the Internet is Hiding from You*. Penguin, London

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Péloquin, T. (Oct. 24, 2020) Une personne sur cinq adhère à des theses complotistes. *La Presse.* https://www.lapresse.ca/covid-19/2020-10-24/sondage-crop/une-personne-sur-cinq-adhere-a-des-theses-complotistes.php

Pew Research Centre (Oct. 17, 2017) *The Shift in the American Public's Political Values*. https://www.pewresearch.org/politics/interactives/political-polarization-1994-2017/

Scherer, M.U. (2015) Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies and Strategies. *Harvard Journal of Law and Technology*. 29:353-400.

Slee, T. "The Incompatible Incentives of Private-Sector AI". In: Dubber, M., Pasquale, F., & Das, S. (2020). *The Oxford Handbook of Ethics of AI*, Oxford University Press.

Silverman, C. (2015) Lies, Damn Lies, and Viral Content. *Tow Center for Digital Journalism*, Columbia University, New York. https://academiccommons.columbia.edu/doi/10.7916/D8Q81RHH

Solon, O. (Nov. 9, 2017) Ex-Facebook president Sean Parker: site made to exploit "human vulnerability". *The Guardian*. https://www.theguardian.com/technology/2017/nov/09/facebook-sean-parker-vulnerability-brain-psychology

Sunstein, C (2007) *Republic.com 2.0*. Princeton, NJ; Oxford: Princeton University Press

Swisher, K (Jan. 25, 2021). The Site Trump Could Run to Next. *New York Times*. https://www.nytimes.com/2021/01/25/opinion/sway-kara-swisher-chris-best.html

Thompson, S. A., & Warzel, C. (2021, January 14). Opinion | How Facebook Incubated the Insurrection. *The New York Times*. https://www.nytimes.com/2021/01/14/opinion/facebook-far-right.html

United States Justice Department. (Feb. 16, 2018) *United States of America vs. Internet Research Agency* (2018) https://www.justice.gov/file/1035477/download

United States House of Representatives (April 11, 2018) *Facebook: Transparency and Use of Consumer Data*. Hearing Before the Committee on Energy and Commerce, House of Representatives (Serial No. 115-144, 115th Congress).

United States Senate (April 10, 2018) *Facebook, Social Media Privacy, And the Use and Abuse of Data*. Joint Hearing Before the Committee on Commerce, Science, and Transportation United States Senate and the Committee on the Judiciary United States Senate. (Serial No. J-115-40. 115th Congress.)

University of Maryland, (2010) Study shows students are addicted to social media. *Communications of the ACM.* https://cacm.acm.org/news/87245-study-shows-students-are-addicted-to-social-media/fulltext

Vaidhyanathan, S. (2018). *Anti-Social Media : How Facebook Disconnects Us and Undermines Democracy*. Oxford University Press.

Yeung, K., Howes, A. and Pogrebna, G. (2020) "AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing" ". In: Dubber, M., Pasquale, F., & Das, S. (2020). *The Oxford Handbook of Ethics of AI*, Oxford University Press.

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

Zuboff, S., (Jan. 29, 2021) The Coup We Are Not Talking About. *The New York Times*. https://www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-society-technology.html?referringSource=articleShare