frontiers in Political Science

ORIGINAL RESEARCH
published: 21 July 2021
doi: 10.3389/fpos.2021.682945

Democracy Under Attack: Challenges of Addressing Ethical Issues of AI and Big Data for More Democratic Digital Media and Societies

Eleni Christodoulou and Kalypso Iordanou**

School of Sciences, University of Central Lancashire, Larnaca, Cyprus

The potency and potential of digital media to contribute to democracy has recently come under intense scrutiny. In the context of rising populism, extremism, digital surveillance and manipulation of data, there has been a shift towards more critical approaches to digital media including its producers and consumers. This shift, concomitant with calls for a path toward digital well-being, warrants a closer investigation into the study of the ethical issues arising from Artificial Intelligence (AI) and Big Data. The use of Big Data and AI in digital media are often incongruent with fundamental democratic principles and human rights. The dominant paradigm is one of covert exploitation, erosion of individual agency and autonomy, and a sheer lack of transparency and accountability, reminiscent of authoritarian dynamics rather than of a digital well-being with equal and active participation of informed citizens. Our paper contributes to the promising research landscape that seeks to address these ethical issues by providing an in-depth analysis of the challenges that stakeholders are faced with when attempts are made to mitigate the negative implications of Big Data and AI. Rich empirical evidence collected from six focus groups, across Europe, with key stakeholders in the area of shaping ethical dimensions of technology, provide useful insights into elucidating the multifaceted dilemmas, tensions and obstacles that stakeholders are confronted with when being tasked to address ethical issues of digital media, with a focus on AI and Big Data. Identifying, discussing and explicating these challenges is a crucial and necessary step if researchers and policymakers are to envisage and design ways and policies to overcome them. Our findings enrich the academic discourse and are useful for practitioners engaging in the pursuit of responsible innovation that protects the well-being of its users while defending the democratic foundations which are at stake.

Keywords: ethics, artificial intelligence, Big Data, digital well-being, digital media, populism, democracy, GDPR

OPEN ACCESS

Edited by:
Leslie Paul Thiele,
University of Florida, United States

Reviewed by:
Elizabeth Shen,
Fudan University, China
John R.T. Bustard,
Ulster University, United Kingdom

****Correspondence:***
Eleni Christodoulou
EChristodoulou3@uclan.ac.uk
Kalypso Iordanou
KIordanou@uclan.ac.uk

Specialty section:
This article was submitted to Politics of Technology, a section of the journal Frontiers in Political Science

Received: *19 March 2021*
Accepted: *09 June 2021*
Published: *21 July 2021*

Citation:
Christodoulou E and Iordanou K (2021) Democracy Under Attack: Challenges of Addressing Ethical Issues of AI and Big Data for More Democratic Digital Media and Societies. Front. Polit. Sci. 3:682945. doi: 10.3389/fpos.2021.682945

> 'Mr. Dorsey and Mr. Zuckerberg's names have never appeared on a ballot. But they have a kind of authority that no elected official on earth can claim.'
>
> (Kevin Roose, New York Times, Jan 9, 2021)

INTRODUCTION

Digital media, by employing Artificial Intelligence (AI) and Big Data, have multifaceted effects on individuals and democratic societies. Algorithms, along with social media, have been identified as two of the major modalities of social influence shaping public opinion in the last decade (Sammut and Bauer, 2021). Before the rise of President Trump in the United States and the Brexit vote in 2016, the discussions around the politics of digital media, often went hand in hand with the awe-inspiring, powerful and emancipatory use of social media. Social media were often seen as both the cause and conduit for strengthening democracy and the rule of law, for inspiring, enabling and accelerating policy changes, institutional reforms, social movements and normative shifts with profound political impacts (Thiele, 2020). From the revolution in Iran (2009), to the Arab Spring in the MENA (Middle East and North Africa) region (2010), the Brazil Spring (2013), the #BlackLivesMatter (2013) and the #MeToo movement in the US (2017), the Gezi Park protests in Turkey (2014) and the Umbrella Movement in Hong Kong (2014), all across the globe, the power of digital media was used to fight social injustices, discrimination, oppression, corruption, police violence and other human rights violations.

The surprising political events of two powerful international actors, the United States and Britain, sowed the seeds for the formation of a more *critical approach* to digital media, one that arguably reached a peak with the 2021 storming of the United States Capitol riots. As a result, we witnessed, possibly for the first time, so-called Big Tech companies like Facebook and Twitter being associated with positions of authority (Roose, 2021) akin to that of a global power; across the news when referring to users' accounts (and in particular those of Trump), we saw words such as "cracked down" (Peters, 2021), "suspended" (BBC News, 2021a), "allowed back" (Hartmans, 2021), "banned" (Stoller and Miller, 2021), "locked" (Rodriguez, 2021) and "pulled the plug" (Manavis, 2021), all verbs associated with power and control. These companies were the decision-makers, while the rest of the world waited to see how they would respond next. These events were significant because thus far, discussions of global governance in political discourse were predominantly associated with intergovernmental organisations like the United Nations, World Bank, International Criminal Court etc., organisations that had legitimate structures with member states and democratic processes.[1] Now, the eyes were on a handful of individuals, self-made, democratically unelected CEOs of these social media platforms. As a *New York Times* commentator put it, Trump's temporary and permanent suspensions from social media clearly illustrated that in the current digital society, power resides "not just in the precedent of law or the checks and balances of government, but in the ability to deny access to the platforms that shape our public discourse" (Roose, 2021).

Undeniably, the digital media debates post-Capitol riots reached a peak, raising concerns about unaccountable and unchecked power that are fundamental to democracies (Jangid et al., 2021).[2] In the aftermath of these events, tech companies were portrayed as "corporate autocracies masquerading as mini-democracies" (Roose, 2021). The Capitol riots raised concerns not only within the United States, but also worldwide. Shortly after the events, the president of the European Commission, Ursula von der Leyen, expressed concerns that "the business model of online platforms has an impact not only on free and fair competition, but also on our democracies" (Amaro, 2021). Despite the significance of these events and their contribution to the shift towards a more critical approach, it is important not to lose sight of the broader context and long-standing threats to democracy. Before the Capitol riots, Big Tech companies were criticised by regulators for their excessive power and for abusing it, for instance, by engaging in illegal tactics in order to stifle competition. In December 2020, the US government and 48 state attorneys general even filed wide-ranging lawsuits against Facebook, claiming that it has a monopoly in the social networking market and that it should be going through divestment (Romm, 2020).[3] Misinformation and fake news; the rise of online radicalisation into Islamist and right-wing extremism; the echo-chambers and political polarization; criticisms regarding the lack of informed consent, autonomy and privacy and heated debates on freedom of speech were all simmering long before the fall of Trump and it is against this wider backdrop that we argue that a fundamental rethinking of the ethics of AI and Big Data is urgently needed and to which our paper contributes. Similar polarised, and polarising debates took place in the context of democratic freedoms being curtailed as a response to the Covid-19 pandemic. Increased digital demands during the pandemic also added weight to the critical approaches towards digital media. Controversy against Google also ensued after the company fired the co-head of its AI ethics unit in February 2021, shortly after another leading researcher in AI

[1]This is not to say that such organisations are not without their democratic weaknesses. Multilateral organisations are highly politicized and it is sometimes the case that countries which contribute more financially or otherwise are regulated to benefit more (e.g., through greater representation). We thank Reviewer one for suggesting this point.

[2]The study by Jangid et al., which analysed over four million tweets and 100,000 parleys found that across both Twitter and Parler, the top ten hashtags between the 7th and 8th of January 2021 were related to the Capitol riots. Also, a search on Google Trends using the term "capitol riot" (geographical region: worldwide) shows how the term went from a value of zero on 5th of January to 100 points by January 7th and was still relatively high at 60 points on January 13th (A value of 100 is the peak popularity for the term and a score of zero means that there was not enough data for the term being searched).

[3]In this way, the lawsuit challenged Facebook's acquisition of Instagram and WhatsApp; regulators asked the court to consider forcing Facebook to sell these off in order to address concerns over competition. We would like to thank Reviewer one for suggesting this example.

ethics claimed she was fired and accused Google of "silencing marginalised voices" (BBC News, 2021b). Both researchers had called for more diversity within the company and were vocal about the negative effects of technology, prompting many to view these departures as an act of censorship on research that took a critical approach on the company's products.

What arguably all the above have in common are signs that the democratic institutions designed to deal with dissent are not working in the current digital age; that the means and modes of negotiating disagreement are neither successful, nor constructive and this adds further urgency to our work. Ignoring people or attempts to censor them are signs of authoritarian governments that do not allow for ambiguity, plurality and tolerance and are likely to do more harm than good in the long-run, including angered users resorting to more niche extremist platforms.[4] If we take the example of Trump, choosing to ban a former president with millions of followers is symbolic not only of the unchecked powers of a handful of unelected entrepreneurs – as the quote at the beginning of the article alludes to – but also of the perilous populist power conferred upon this individual through digital tools, the power to influence the "hearts and minds" and behaviour of hundreds of millions of people. At the same time, its success as a tool for "giving voice", is an illustration of the weaknesses of the often distant, bureaucratic and staged communication of representative democracy, and the thirst for more tools of *direct* democracy – something Trump picked up on early on.

Recent studies have identified and discussed ethical issues that arise from AI and Big Data either on a normative or empirical level (Jobin et al., 2019; Müller, 2020; Ryan et al., 2021; Stahl et al., 2021). What is missing from existing literature however, is a comprehensive understanding of and engagement with, the *challenges* that arise when such attempts are made to mitigate the risks of Big Data and AI while ensuring that they are harnessed for the benefit and well-being of society. Existing studies also rarely include focus groups directly from the primary actors – the key stakeholders themselves. Our paper contributes to the research landscape of addressing ethical issues by bringing to the fore the multifaceted dilemmas, struggles, tensions and obstacles that stakeholders are confronted with when being tasked to present solutions to ethical problems of Big Data and AI. Stakeholders included policymakers, NGO representatives, banking sector employees, interdisciplinary researchers/academics and engineers specialising in AI and Big Data. Our main research question is: *What are the core challenges that stakeholders face when addressing ethical issues of AI and Big Data?* Identifying, discussing and explicating these challenges is a necessary step if researchers and policymakers are to envisage and design ways and policies to overcome them. This is an urgent task given the need for a renewal of democratic values, for more power checks and controls and for the prevention of further human rights violations.

The symbiotic relationship that populist leaders often have with digital media (Postill, 2018; Schroeder, 2018), with both of them thriving (as one reinforces the popularity/profitability of the other), raises fundamental questions regarding the *compatibility* of ethical and healthy democracies with the business models of digital media companies. This is a further sign that we can no longer ignore the need to address ethical issues of AI and Big Data. It is crucial to do so if we are not to repeat the mistakes of the past that led to digital media being used to violate democratic principles, abuse human rights, spread panic, misinformation and fake news as well as extremist propaganda. It is within this specific context of moving towards an *ethical path of digital well-being* that we locate our work. According to Burr and Floridi (2020), digital well-being can be defined as "the project of studying the impact that digital technologies, such as social media, smartphones, and AI, have had on our well-being and our self-understanding" (p. 3). In our work we focus on the ethical implications of digital well-being vis-à-vis the protection of human rights and the preservation of fundamental democratic principles.

Two premises underpin our work. Firstly, that the damage done so far is neither irrevocable (Burr and Floridi, 2020), nor inevitable. Secondly, we concur with Müller (2020) who argues that we should avoid treating ethical issues as though they are fixed: we neither know exactly what the future of technology will bring, nor do we have a definitive answer as to which are the most ethical practices of AI and Big Data let alone how to achieve these. Therefore, "for a problem to qualify as a problem for AI ethics would require that we do *not* readily know what the right thing to do is" (Müller, 2020). This does not preclude us from adopting certain approaches as more ethical than others; it merely highlights that in that case it is no longer an ethical dilemma.

To examine stakeholders' views, six focus groups were organized across Europe and analysed taking a data-driven approach in order to best reflect stakeholders' concerns and experiences. Rather than relying on desk research, our rigorous empirical approach sought to find answers directly from the primary actors involved in these processes, generating novel and useful insights from stakeholders in the area of shaping the ethical dimensions of technology use. The findings presented in this paper therefore make a significant contribution to the ethical discussion of Big Data and AI by introducing a conceptual apparatus that increases our knowledge of the issues that are at stake when addressing ethical issues and the complexity they involve. Our findings enrich the academic discourse while also providing useful insights and suggestions to policymakers and organisations that are engaging in the pursuit of responsible innovation that protects the well-being of its users.

DIGITAL MEDIA, BIG DATA, ARTIFICIAL INTELLIGENCE AND DEMOCRACY

In the current section, we focus on the threats that digital media impose on democracy with a particular focus on Big Data and AI. We adopt a broad understanding of democracy that includes both

[4]As Simon Goldhill put it "it is not by chance that Plato was the favourite philosopher of both Hitler and Stalin" (2021). Simon Goldhill, "Power and Impunity: what Donald Trump and Boris didn't learn from the Ancient Greeks", LSE Public Lectures and Events, 29 January 2021, https://www.lse.ac.uk/lse-player?id=da1c939f-2756-4018-81d6-381413baaf31

direct and indirect democracy rather than focusing on specific geopolitical contexts or particular democratic systems. Principles of democracy that underpin all types of democracy include human rights, justice, freedom, legitimacy and checks and balances.

A tremendous amount of digital trace data – Big Data – is collected behind the screens using the trail we leave behind us while we navigate and interact with digital media, through clicks, tweets, likes, GPS coordinates, timestamps (Lewis and Westlund, 2015), and sensors by smart information systems, which register information of our behaviour, beliefs and preferences for profiling citizens. Those who have access to citizens' digital data and profiling know more about individuals than probably their friends, family members or even individuals themselves (Smolan, 2016). That data is then used to tailor the information we receive in an attempt to influence our behaviour for the sponsors' benefit, for securing, for example, financial profit or winning the elections, as was the case with Cambridge Analytica (Isaak and Hanna, 2018). Techniques that were initially developed in the sphere of marketing and advertising, employed by sellers, including exploitation of behavioural biases, deception, and addiction generation to maximize profit (Costa and Halpern, 2019), are now used in the sphere of politics, for instance to manipulate public opinion and maximize votes (Woolley and Howard, 2016). Algorithms, by using users' profiles based on their previous interactions online, can provide the kind of input that is more likely to influence a particular individual (Müller, 2020). In times of elections in particular, those who control digital media have the potential to "nudge" or influence undecided voters and win elections, leading to a new form of dictatorship (Helbing et al., 2019; Roose, 2021) and traumatizing democracy.

Besides the orchestrating efforts of nudging (Helbing et al., 2019), digital media constitutes a fertile ground for the spread of misinformation through the profound absence of any form of gatekeeping. Misinformation, hate speech and conspiracy theories have found a way to reach thousands of citizens through digital media, especially social media, threatening political and social stability (Frank, 2021). Although concerns regarding the deliberate dissemination of the information in order to affect public perception were evident before (Bauer and Gregory, 2007), those issues have been amplified with the use of AI in digital media. Misinformation can be disseminated by different actors, including politicians, news media and ordinary citizens (Hameleers et al., 2020), but also machines, such as the Russian propaganda bots that infiltrated Twitter and Facebook (Scheufele and Krause, 2019).

Digital media, besides using deliberate processes, such as altering evidence and purposefully fuelling fake news, employ other mechanisms which do not contribute to the protection of human rights – a fundamental element of a democratic society. Algorithmic filtering, which refers to prioritizing the selection, sequence and the visibility of posts (Bucher, 2012), and is embedded in online social platforms, reinforces individuals' pre-existing beliefs and worldviews (Loader et al., 2016; Gillespie, 2018; Talmud and Mesch, 2020), increasing biases as well as social and political polarization (Helbing et al., 2019). Given the biases in favour of one's own position and the limited critical evaluation of evidence reported when individuals are reading new information (Iordanou et al., 2020; Iordanou et al., 2019), restricting one's input of information to only that which is in alignment with one's beliefs, impedes self-reflection (Iordanou and Kuhn, 2020) and contributes to polarization and extremism. A study by Ali et al. (2019) found that the algorithms behind Facebook's delivery of political advertisements disseminates ads using the criterion of alignment between the inferred political profile of the user with the advertised content, "inhibiting political campaigns" ability to reach voters with diverse political views. These findings provide evidence of how social media algorithms contribute to political polarization. This is very concerning in light of findings showing that interaction with individuals who share different views from one's own are vital for the development of critical thinking (Iordanou and Kuhn, 2020). Social media platforms started and continue to run as business models, aiming to generate revenue by directing ads to users based on their digital profile (Frank, 2021). Investing on individuals' need to socialize, especially in the absence of profound alternatives, and utilizing human minds' vulnerabilities – e.g. proneness to addiction – digital media have undertaken roles that they have not been designed for, such as being the main medium that citizens use to read news and get information on important issues of their personal and social life. More than eight-in-ten Americans get news from digital devices (Shearer, 2021). Having not been designed to inform or educate, and in the absence of a regulatory framework or any other mechanism of checking the role of digital media as information providers or asking them to be accountable for their actions (Cave, 2019), findings showing a link between social media use and lower levels of political knowledge (Cacciatore et al., 2018) are not surprising. Getting news from social media was found to be related with uncivil discussions and unfriending, that is shutting down disagreeing voices, contributing to polarization (Goyanes et al., 2021).

Another concerning issue to democracy reported in the literature, is the exercise by digital media of power structures and biases in society (Diakopoulos, 2015: 398). The digital media may encapsulate the worldviews and biases of their creators (Broussard, 2018; Noble, 2018) or the data they rely on (Cave, 2019). This effect can be especially detrimental for adolescents, for which digital media and cyberspace is an integral part of their social life, and who are in a critical stage of development and socialization (Talmud and Mesch, 2020). The result of algorithmic bias or bias in Big Data is the replication of biases, discriminatory decisions and undemocratic situations (Pols and Spahn, 2015). One example is the scandal with Amazon's recruitment AI tool that was scrapped after it was revealed that it was discriminating against women (Cave, 2019). There is certainly a need to make algorithms, AI, and digital media more fair, transparent, and accountable (Wachter et al., 2017). The ethical implications of AI, Big Data and digital media on democracy form the backdrop of our work and set the scene for better understanding the challenges stakeholders face when trying to mitigate or prevent these ethical implications.

TABLE 1 | Main Ethical Issues regarding AI and Big Data, identified by stakeholders in Iordanou et al. (2020).

1	Loss of autonomy, human decision-making and control: "A very, very dangerous direction"
2	Loss of privacy: monitoring and surveillance
3	Prioritisation of financial over ethical interests: Big Tech companies manipulating users
4	Lack of (access to) information and knowledge
5	Biased, inaccurate data and algorithmic bias
6	Human jobs replaced by machines and dangers of machines and apps
7	Loss of access to services
8	Loss of trust
9	Lack of accountability
10	Threats to (and violations of) human rights

THE PRESENT WORK

In the present work, we examine different stakeholders' views on what they consider as the major challenges that we need to address to deal with ethical issues related with AI and Big Data. The present work is part of SHERPA, a Horizon 2020 project funded by the European Commission. Building on the ethical issues that have been identified by stakeholders as the major ethical issues arising from AI and Big Data, in earlier work (see **Table 1** and Iordanou et al., 2020), in the present work we focus on stakeholders' views regarding what they consider as the *challenges* that need to be alleviated for making AI and Big Data more ethical and aligned with human rights.

METHODOLOGY

Data Collection

Focus groups with stakeholders were chosen as an appropriate research method as they provide a constructive and close-knit environment for stimulating expert discussions on niche topics. They are also useful for collecting diverse expertise as well as showcasing "consensus (or dissensus)" in ways that a one to one interview may not (Pierce, 2008). This was particularly beneficial for the present study given that addressing the challenges of Big Data and AI is still at an embryonic stage. Focus groups are also ideal for examining social groups' views, in this case their social representation of technology (Bauer and Gaskell, 1999).

Although there was some degree of flexibility in the focus group discussions, the focus groups followed the same questions, which revolved around three core aspects: a) the main ethical issues that come out of AI and Big Data and their relation to human rights (see **Table 1**); b) the nature and limitations of current efforts to address these ethical issues; c) and suggestions of activities that should be undertaken in the future to deal with ethical issues that have not yet been adequately addressed. The questions posed to the participants did not offer specific options but rather were open-ended in order for the views of the participants to emerge rather than for the facilitator of the focus group to impose their own views or influence the discussion (see the Appendix for a list of the questions). In the current paper, we focus on the challenges reported by the participants which emerged mostly from the answers given to part (b) i.e. the various obstacles, problems, tensions, difficulties and dilemmas stakeholders tend to encounter when attempting to tackle ethical issues related to AI and Big Data. Often times this took a form of explanation as to why there has not been adequate progress so far. These challenges were also sometimes expressed as limitations of current efforts or gaps identified in current efforts.

Participants were sixty-three individuals who participated in six focus groups. Experts on AI and/or Big Data were recruited, either through their participation in relevant conferences or through personal networks of partners of the project SHERPA. Expert voices included policymakers, NGO representatives, banking sector employees, interdisciplinary researchers/academics and engineers specialising in AI and Big Data. In particular, two focus groups, (n = 7 and n = 8) took place in the context of the ETHICOMP 2020 conference (International Conference on the Ethical and Social issues in Information and Communication Technologies). A third focus group (n = 12) took place in the context of the United Kingdom Academy for Information Systems' (UKAIS) conference, which is attended by researchers and practitioners of information systems. The fourth focus group (n = 19) took place with SHERPA Stakeholder Board members. For the last two focus groups (n = 8 and n = 9) the organizers recruited participants by reaching out to them from their contact list, based on their expertise.

The present study was conducted with ethics approval from the Cyprus National Bioethics Committee. Prior to the focus groups, the participants were provided with an information sheet and their written consent was secured. One focus group took place in the participants' native language, Greek, and was translated in English. The rest of the focus groups were conducted in English. The duration of the focus groups was on average between 60 and 90 min. The focus groups were recorded and transcribed and the transcripts were anonymised. One of the focus groups – pursued in January 2020 – was conducted face-to-face in Cyprus, while the other five – pursued between March to May 2020 – were pursued virtually given the restrictions imposed by the Covid-19 pandemic. Three individuals chaired the focus groups (one individual chaired three focus groups, another individual chaired two focus groups, and a third individual chaired one focus group). All the chairs participated in a training workshop, organized by one of the authors, to ensure consistency in the data collection among the different focus groups.

Data Analysis

The data obtained during the focus groups was analysed using thematic analysis and in particular the framework provided by Braun and Clarke (2006). We followed the six stages of thematic analysis (2006, p. 87): 1) Initial data familiarisation; 2) Generation of initial codes; 3) Search for themes; 4) Review of themes in relation to coded extracts; 5) Definition and final naming of themes; 6) Production of the scholarly report. Thematic analysis can be defined as: "a method for identifying, analysing and

TABLE 2 | Eight themes resulting from the focus group analysis.

1	Difficulties in reconciling heterogeneous perspectives to ethics
2	The rise of polarisation and populism and the weakening of democracy
3	The engineers' lack of self-awareness and critical self-reflection and the difficulty in educating them
4	Not enough pressure on AI developers/producers to address ethical issues: "no punishment for the bad actors"
5	Ethics washing and the "checklist" approach
6	Users themselves choosing convenience over privacy/other ethical concerns
7	Challenges pertaining to legislation and regulation: the case of GDPR
8	Health crises: the case of the Covid-19 pandemic

reporting patterns (themes) within data. It minimally organizes and describes [the] data set in (rich) detail" (Braun and Clarke, 2006, p. 79). Although thematic analysis always involves a degree of interpretation (Boyatzis, 1998), due to the main objective of the focus groups, which was to let the participants' opinions and expertise inform the findings, we took a data driven, inductive approach in which codes and themes were generated from the data (open coding) rather than having a codebook prepared in advance of the data analysis (deductive approach). This meant also that themes were identified at what has been called a "semantic" level; in contrast to a thematic or discourse analysis at the "latent" level where the researcher is trying to uncover hidden assumptions and ideologies and unmask power asymmetries, our goal in this study was to use the experiences and insights provided as experts' answers to the posed questions. Therefore the themes that emerged were "identified within the explicit or surface meanings of the data" (Braun and Clarke, 2006, p. 84).

Given the number of focus groups analysed and to ensure better organisation, classification and interpretation of the data, the analysis was performed using Nvivo, a qualitative data analysis computer software (Version 12). One of the authors was involved in the data analysis for stages 1–3 in order to ensure consistency across coding of the data. After this the researcher shared the resulting codebook with the other author and discussions and brainstorming led to completion of stages 4–6 of thematic analysis in close collaboration with each other. These included weekly meetings to discuss the hierarchy of codes, the meaning of each code, resolve any dilemmas and "as a means of reflexively improving the analysis by provoking dialogue between researchers" (O'Connor and Joffe, 2020, p. 6). Beyond the reflexive benefits, thorough shared documentation, internal audits and regular meetings acted as "an incentive to maintain high coding standards" on an intracoder level (White et al., 2012; O'Connor and Joffe, 2020; p. 4). The final data analysis for the specific topic of "challenges" presented in this paper resulted in a total of 71 codes (of different hierarchies, "parent", "child", "grandchild") which were then collated into a total of eight themes (see **Table 2**). For the coding process, we followed closely the instructions and advice of Braun and Clarke (2006), Bryman (2008) and Charmaz (2004), namely to code as thoroughly as possible – line by line so as not to lose any detail or potential interpretation of the data – and not to try and smoothen out contradictions, inconsistencies or disagreements as these were an inevitable and often useful part of the data.

RESULTS

Challenges Emerging When Addressing Ethical Issues

The data analysis identified eight core themes (challenges) that stakeholders are faced with when attempting to address ethical issues related to Big Data and AI. These are outlined in **Table 2** (below). Before presenting each theme in detail, it is important to note some background information that was provided by the participants that form the backdrop against which attempts to address ethical issues take place and which help further our understanding of the complexity involved and of the themes that emerge.

Firstly, and perhaps unsurprisingly given the digital demands of the pandemic, the ubiquity and inevitability of technology was noted. It had now become a dominant, inescapable part of our lives and "opting out" of it while remaining an active part of society had become virtually impossible: "I don't know if there is any turning back at this point in the level of technology that we all depend on so much" (Focus Group B) noted one individual while another sought to emphasise how it has become enmeshed into our everyday experiences and influences us across different stages of our lives: "It is deeply embedded in our lives . . . Google uses an algorithm in order to provide us with information and we all use Google in our daily lives to find information from the age of 5" (Focus Group F). Associated with this was the ethical issue of surveillance as search engines, digital devices and digital media, seen as "tracking us constantly" (Focus Group A). The second piece of contextual information was related to the complex and fast-paced nature of technology. Participants admitted the complexity of AI and Big Data, yet pointed out how this complexity is sometimes used as an "excuse" to absolve companies from the "moral consequences of the[ir] decision":

> 'sometimes the black box is used as a good excuse by some organisations saying, "This is just a deep-learning system. We don't know what is happening but the results are great"' (Focus Group A).

The speed of technology was contrasted with the slow pace of policy formulation and implementation: "the technology advances and then the policies start to follow" (Focus Group A). Technology was viewed as growing faster than the ability to legislate in order to set some ethical constraints to its misuse and abuse. Regarding such initiatives, these were sometimes discussed in general terms, while at other times reference to specific regulation, legislation or codes of conduct were made, for instance such as the Association for Computing Machinery's

(ACM) updated Code of Ethics and Professional Conduct (2018).[5] The legislation (one of the eight themes) most discussed was that of GDPR (General Data Protection Regulation), a regulation in European Union (EU) law on data protection and privacy in the EU, passed in 2016 and implemented in 2018 (see Danidou, 2020). During the focus group discussions, some saw current efforts as a glass half empty, focusing on the limited "range of tools" to reduce the ethical issues that arise (Focus Group A) and their weaknesses while others saw the glass half full, noting the progress that has been made compared to previous years.

Difficulties in Reconciling Heterogeneous Perspectives in Defining and Addressing Ethics

Several participants across the focus groups noted that one key challenge was agreeing on what constitutes an ethical framework in the first place, given the diverse and sometimes conflicting perspectives and assumptions on this topic. "We are speaking a different language" noted one participant when referring to discussions they had as an academic and technology specialist with legislators (Focus Group B). Indeed, across the focus groups the type of profession which affected people's understanding and perception of the topic, for instance an academic vs. a company executive or a technology expert vs. a legislator, were seen as variables creating differences and sometimes irreconcilable tensions in terms of how to best tackle ethical issues.

Various other factors and considerations were mentioned that affected people's interpretations and how seriously they viewed ethical issues. Some sources of heterogeneity were identified as emerging from factors such as different generations and age groups; gender; whether it was a developer or a consumer:

> '[attention paid to ethical issues] varies in terms of the perception of people who work with it [technology], as well as the perception of the people who use it. You might find great differences between, or major differences between individuals, based on their age group, or even based on their gender as well … the interpretation of the ethical issues is highly subjective between different human beings' (Focus Group A).

> 'I do hear, especially from young people, when you try to tell them about Facebook or any other social media application, when they post their private information and things like that, and when you explain to them you get the answer "I don't care. It's okay for people to see my information". And when you explain that it's dangerous, they say "It's fine, it's okay". So, the mentality of the young generation it's very different' (Focus Group F).

Additional variables were seen as the size of a company; the position of one in the company (for instance, a CEO vs. an employee); and the academic discipline (for instance philosophy vs. computer science):

> 'How does a small SME cope with this kind of issue, compared to a big company? How does a person who's not got much, perhaps, educational understanding or many finances deal with this, compared to someone who doesn't?' (Focus Group A).

> 'It's the perception of the ethical issues, or perceptions of data protection that can be very different between people who are working on the same project. Also, the role that those people have. Sometimes, the kind of power that a permanent member of a staff within an organisation can be different from temporary staff, a consultant, so that is also important-power dynamics within an organisation' (Focus Group A).

> 'So, I tried to go around as someone that is concerned about privacy in general to try to break the walls that exist between the different areas … this is a very diverse group of people here but we all live in our own worlds and we only speak with our colleagues and our peers and that's it' (Focus Group B).

A final variable that increased heterogeneity was related to regional, cultural and country differences. For instance, one participant emphasised how 'What is ethical in one culture is not necessarily ethical in another culture.' (Focus Group A). Regional and country differences were also noted:

> 'There's a difference between Northern Europe and Southern Europe. If that puzzles you, I'm not surprised, but I do see a different kind of society in Japan from the UK, although we have strong analogies' (Focus Group A).

Ultimately, the fact that "everyone's definition of what is ethical or where to draw the line will differ" complicated an already multifaceted problem, added further tensions and meant that paving a common way forward became almost unmanageable.

The Rise of Polarisation and Populism and the Weakening of Democracy

A second challenge that focus group participants identified was the rise of nationalism and populism and a weakening of democratic institutions. The underlying issue at stake was how can one even begin tackling ethical issues when the very foundations upon which democracy is built are being undermined. Human rights violations and unethical ways of behaviour were seen as going hand in hand. The weakening of democratic institutions through the rise of populist politicians across the globe, an increase in nationalism, undemocratic behaviours by the governments and unregulated fake news and hate speech were all seen as challenges that hindered

[5]Reflecting the increasing pervasiveness of computer systems in society, the ACM revised its Code in 2018, more than 2 decades since it was last updated in 1992. The Code is "is a collection of principles and guidelines designed to help computing professionals make ethically responsible decisions in professional practice. It translates broad ethical principles into concrete statements about professional conduct" (ACM, 2018).

attempts to address ethical issues and human right violations. There were various references to both specific scandals such as Cambridge Analytica but also country contexts where human rights were viewed as under threat or violated, with the public having little agency to resist or reverse the negative consequences. For instance, government surveillance and monitoring, data collection without informed and adequate consent, and manipulation of data for either financial or political gain, were seen as directly related to human rights violations. Some participants were quick to point out that human rights violations and weak democracies are a characteristic of even the more "established" or "progressive" democracies such as the US (with references to Trump) and the United Kingdom (with references to Boris). More explicitly authoritarian practices were discussed in the context of China and Brazil. It was said that China could "switch you out of society" as a result of monitoring individuals' social media accounts (Focus Group A). Another example was given in relation to Brazil with one participant who was based there, noting that the current president is acting like "a dictator", suspending all legislation related to data protection, privacy and freedom of information (Focus Group B):

> 'at the moment we have a president that is kind of a dictator and recently we had a set of laws, they are very, very new in comparison with the situation in the UK, in the US or in Europe, but we have freedom of information, we have privacy protection act, data protection act and very recently he suspended the use of all those specific laws because he was involved in a very bad scandal of you know fake news and robots and stuff' (Focus Group B).

Participants spoke of populist leaders and governments (which sometimes cooperated with private companies) intentionally and actively denying people data privacy and anonymity in such a way that the public would often not even identify these actions as unethical and therefore as something that needed to be addressed. As one participant put it:

> 'when we leave giant companies gathering data and when these data is used to influence the population, we can have a population that sees a leader suspending this kind of law and doesn't think that would be a problem because they think that in the same way that companies and people that benefit from it like us to think, that if I have nothing to hide why would I not want to be surveilled' (Focus Group B).

One participant spoke about a seemingly disintegrating world were basic institutions such as the World Health Organisation or the European Union were being challenged. In this "dangerous" context of political disintegration and polarisation, politicians were often resorting to blame games rather than addressing the ethical issues head-on and protecting human rights:

> 'We see from President Trump's take on life that he wants to, effectively, fragment the World Health Organisation as we speak; the British perspective to fragment Europe, which is going on in association with that. While the United Nations is out there, as well, charities are in meltdown. Therefore, this is a very critical time to retain some semblance of regulation and human rights on a global scale, due to the rise of nationalism, really, and blame games that are going on (Focus Group A).'

Often the climate of fake news, polarisation and hate speech in digital media—accelerated by algorithms[6] - presented a fertile ground for populist leaders to grow their impact and influence. It was therefore, particularly challenging to expect any support for more ethical policies, let alone proactive initiatives, by state leaders who thrived and depended on the reproduction of such dynamics in society, but also sometimes actively "work [ed] towards" creating them in the first place (Focus Group A).

The Engineers' Lack of Self-Awareness and Critical Self-Reflection and the Difficulty in Educating Them

The third challenge identified was related to the software engineer/developer and what was a perceived lack of ethical self-reflection and critical thinking in order to be able to recognise their own biases and individual responsibility. Some participants placed responsibility on the developers, others on the companies while others appreciated that both had their limits and that there were other factors, some outside their control, that hindered further progress in making technologies more ethical.

One pitfall presented here was a simplistic binary of good vs. evil which made it difficult for developers to be ethically self-reflective and pro-actively engage with ethical issues so as to prevent embedding discriminatory practices such as racism and sexism in algorithms. As one participant pointed out:

> 'most of us think we're ethical and we operate with a very bad ethical premise that says I'm a good person and evil is caused by evil people. I'm not an evil person. So I don't have to worry about it. So when I write the algorithm, I'm a good software engineer. I don't even have to question this. I'm doing a fine job' (Focus Group C).

A related point was made in terms of being aware of one's own subjective cultural norms that may affect one's decisions and designs. Not only was this critical awareness often missing, but crucially, the nature of software is such that it is difficult to change these underlying norms once the system is built:

[6]Some participants argued that digital media and AI accelerate and amplify the spread of fake news: AI systems used in Facebook and Google make this problem even bigger because if someone starts a rumour it spreads faster and to many more people because of AI. Because it will contextualize . . . It amplifies because of all the interaction with the fake news. Let's say now with the Coronavirus, people are sharing and sharing and sharing and it just spreads' (Focus Group E).

> 'the cultural norms that we have, but don't even realise we have, that we use in order to make decisions about what's right and wrong in context. It's very difficult for any software system, even a really... advanced one, to transcend its current context. It's locked in to however it was framed, in whatever social norms were in place amongst the developers at the time it was built' (Focus Group C).

Another concern raised regarding AI developers by some participants was that they usually "represent only a niche of society, a particular niche society" and they do not always have the required pluralistic, diverse and broad perspective so as to build 'inclusive technologies' (Focus Group C). Agreeing with this point one participant who referred to himself as 'an old white guy' seconded this opinion arguing that narrow viewpoints of developers extend into and are reflected in the software:

> 'I'm niche market and I do the photo recognition software and I'm an old white guy. So the only people I recognise are white males with beards. And that happens in the software, we know it's happened and we've framed out the ethics' (Focus Group C).

Therefore, this specific challenge can be seen as related not only to the practices of the employees at a particular company but also to the hiring practices of the company itself which did not ensure that their team was diverse enough or educated adequately on ethical practices.

The education of companies in relation to ethical issues was seen as a challenging task, not just in terms of lacking financial incentives but also ethical incentives or lack of an 'organization culture' of ethics in a company: 'You need to convince the managers' remarked one participant who agreed with another participant who suggested that such issues may not 'affect them in any way. So, they don't have to care about it. That's why it's harder for them to apply it anyway' (Focus Group E).

A more practical obstacle that was highlighted by some participants was the lack of guidance and the difficulty of practically educating designers due to the complexity and unpredictability of technology. Even when designers have the required will to make ethically suitable designs, it was argued that it is often difficult to provide them with 'concrete guidance' due to the complex nature of Big Data and AI (Focus Group B). In terms of unpredictability, there is a lack of 'work looking at scenarios of unintended consequences' precisely because 'we don't know the unintended consequences of the decision-making of the machine' (Focus Group A).

Others added that even when ethical courses did exist, the way the ethical issues were communicated was a user-friendly one, neither in terms of the user interface nor the user experience (Focus Group F). A final difficulty presented in relation to educating developers was that whereas education was often seen as an individual process, 'algorithms generally are used by companies' and so this brought up the task of education at a more collective, company level that was hard to achieve (Focus Group F).

Not Enough Pressure on AI Developers/Producers to Address Ethical Issues: "No Punishment for the Bad Actors"

The difficulty of educating individuals or companies leads to the fourth challenge identified when trying to address ethical issues of AI and Big Data: the lack of pressure on AI developers to take responsibility and adequately address ethical issues. It was not always clear who should exert this pressure, though some mentioned in their suggestions the need for effective sanctions for violating regulations and potential avenues to improve accountability such as further legislative frameworks and public pressure.

One participant who said they had substantial experience with producers of AI-based technologies and solutions stated that because the latter are not really interested in or motivated to address these issues - especially when this would increase costs—their approach was one that tried to ascertain 'what is the minimum we have to make to be according to the law and not to address the issues really in full' (Focus Group B). This approach was along similar lines to what one other participant called 'a checklist approach', merely to be able to tick the legal boxes in a superficial way that ensured the companies were allowed to operate by law even if the ethical issues were essentially left unaddressed or under-addressed (see also the fifth theme below on 'ethics washing'). Some participants agreed that it is ultimately the responsibility of the company manager, but they are the hardest to convince 'because they don't have it as a priority and 'it costs money without a direct effect' (Focus Group E).

Big Tech companies like Facebook were mentioned as an example of how companies manage to 'get away with things' (Focus Group A) when malpractice has occurred, despite laws and regulations and this indicated a strong limitation or even failure of existing efforts to address ethical issues. This is related to the theme of the 'lack of accountability' that emerged as one of the main ethical issues of the first part of this study (see methodology section above). As one participant put it, it often seems like 'there is no punishment for the bad actors', no deterrent to prevent them from unethical practices (Focus Group C):

> 'Accountability is the key that is not adequately addressed yet. We have Cambridge Analytica, but the Chief Executive didn't go to prison. We have other people who are actually manipulating data for political and commercial reasons, but nothing happens. They get fined by a miniscule amount of money, so, therefore, accountability is not adequate.' (Focus Group A).

Another reason provided for the lack of pressure was the fact that Big Tech have managed to have 'minimum regulatory intrusion' because they leverage their financial and political power to successfully 'lobby the legislators that are supposed to be regulating them' in the first place (Focus Group B). This also relates to the second theme above, where the governments and parliaments that are democratically elected to protect the public

end up promoting the vested interests of private companies instead.

Ethics Washing, Empty Promises and the 'Checklist' Approach

When attempts were made to address ethical issues pertaining to Big Data and AI, superficial approaches were seen as an obstacle to genuine transformation and progress. Some IT companies were seen as giving empty promises or overpromising but not delivering (Focus Group A). Other companies tended to present the final end, a seemingly positive end point, as a means to justify unethical means, thereby absolving themselves of the 'moral consequences' of their decisions; such projects were "misleading" and 'create [d] unfairness' (Focus Group A). Therefore, it was argued, this involved intentionally altering perceptions through the use of deception.

Participants also referred to 'ethics washing', that certain large corporations merely want to give the impression that they 'are paying attention to ethics, developing ethics boards and so forth' because they have a product to sell and if it looks ethical or they say it is ethical, that will help their sales, even when it is not 'actually better in ethical terms' (Focus Group B and C).

One participant expressed strong concerns about what he referred to as the 'checklist' approach to ethical issues which he argued presented a hindrance to progress (Focus Group C). He was critical of the AI community in its approach to ethics, arguing that it 'thinks it is inventing ethics' and that organisations writing ethical standards are currently doing so without looking at previous efforts in other areas of ethics. They are therefore lacking context and not trying to learn from past mistakes. Explaining what he meant by the 'checklist approach' he criticised what he saw as a very mechanical, superficial way of approaching ethical issues:

> 'they're producing a standard, a checklist, a thing that you do as if, "I do this, this, this and this, my AI will be OK"...if you have a compliance checklist, what happens, at least in companies, is that checking the box is the consideration rather than the ethical impact of what you're doing. So did I conduct a test?..So I get to check this box and I'm done, not a question of how it impacts others or raising other kinds of questions, but just did I do this kind of test? Yes. Have I got a comment in the code? Yes. And it's not a question about its ethical impact . . . And if you do this, you're doing good AI. So did you test that you coded properly that your programme doesn't crash? Yes, I did. Did you check that if people try and use it, they'll move their hand too fast and will get carpal tunnel syndrome? Well, no, that's an ethical issue. I don't have to do that and I don't have to deal with this' (Focus Group C).

The participant was also critical of the language used i.e. 'codes' of ethics, which were treated as checklists, rather than explaining why certain values are important and why programmers should care and deal with these aspects. This he argued is a limitation as ethical codes are currently being 'treated as constraints rather than opportunities for goodness'. In other words, they are not used in a constructive way but as something that people fear they need to comply with or else face repercussions. An exception to this, he argued, was the ACM Code of Ethics which instead of constraining the way that computing professionals could operate, focused on opportunities and responsibilities for improving society and working with stakeholders.

Finally, the participant argued that this 'checklist approach' is a limitation that can be found in recent EU regulations/codes of ethics that were released in late 2019. Again, this approach, it was argued, focused on producing quality software rather than on how to best support and improve society and stakeholders.

Users Themselves Choosing Convenience Over Privacy/Other Ethical Concerns

The sixth challenge was related to the end-user putting access to digital media and digital services and quality of these services above ethical issues. The core argument here was that even when users to some extent knew about certain data collection breaching privacy, some still chose access to a service and getting 'the job done' rather than paying attention to the ethical issues at stake. A prominent rationale given was that people 'love technology' (Focus Group B) and tend to avoid taking serious action in response to ethical concerns 'until something bad happens to you, personally, or on a larger scale' (Focus Group A).

One participant referred to a research they were involved in which found differences between Generation Z and Millennials[7]:

> 'we found that those younger consumers or individuals who come who are part of Generation Z are actually, sort of, okay with a trade between privacy and personalisation. They pay less attention to these ethical issues . . . as long as they have a service delivered to them, the required quality and at the same time the job is getting done . . . but when it came to Millennials . . . things changed...They completely stopped using the system . . . the trust issues were a major thing for them' (Focus Group A).

Giving the example of smart home devices such as Alexa or Siri, one participant remarked that when having the dilemma of convenience vs. privacy or security-for instance, having the application to be constantly listening to your discussions so that it responds when you call it vs. having to press a button to activate it—then '[a]lmost every user chooses the convenience over privacy' (Focus Group E). A similar remark was made in relation to digital media:

[7]Millennials tends to refer to those born between 1981 and 1996. Generation Z refers to those born from 1997 onward. For further information see here: https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/

> 'I still see Facebook and Instagram and Twitter and WhatsApp and Zoom for instance gathering data and leaving back doors in a computer science point of view that can gather data and people are still being happy to use all those applications' (Focus Group B).

Many users, one participant argued, even viewed the possibilities that emerged with data collection as 'a gift' and so failed to consider it as an ethical issue:

> 'the next restaurant I'm going to is suggested by Google because it collects information about where I go about, and where next time I should go and we always take it as a gift, that's alright' (Focus Group B).

A prevalent discourse was that the responsibility was partly of the user, but *also* partly of the company given the lack of (accessible) information digital media companies provided to the users in order for them to be able to make informed decisions and choices. Regarding the responsibility of the user, there was disagreement as some participants believed there is 'public awareness' but what is lacking is 'the will to do anything serious about it'; 'You just think … "Well it's not an issue until I have to deal with it"' (Focus Group A). Others disagreed and argued that:

> 'customers awareness is really low regarding those issues, so they are not demanding from the producers, protecting their rights and addressing those ethical issues so the producers don't' (Focus Group B).

Regarding the actions of the users being interlinked with the responsibility of the company/developers, there were several discussions that highlighted not only the responsibility of the company/developers but also of the limited agency that the end-user had ultimately. Firstly, it was highlighted that user's options and choices seem like choices but in practice these are dilemmas that are not easily resolved (see also discussion at the beginning of this section on the difficulty of choosing to 'opt-out' from access to digital media). As one participant emphasised:

> 'I think we need to bear in mind that [in] a lot of ethical issues you have like a right of conscience that you can opt out, or you can take objection to something. I think it's becoming increasingly difficult in this area, and we should be aware of that' (Focus Group A).

Secondly, companies also often downplayed the negative implications which prevented the public from being truly aware of the extent of the malpractice:

> 'tech giants tend to tell us that we shouldn't worry about surveillance. That if we're not doing anything wrong, you know, you have nothing to hide then what's the problem and part of the problem is democracy and expanding democratic rights, whether it's the civil rights of people of colour or if it's women or the environment now … democratic citizens have a right to … privacy and if that right is compromised it's not simply your own free will that's at stake. It's the entire range of human rights, democratic rights such as equality, freedom of expression, you name it' (Focus Group B).

Challenges Pertaining to Legislation and Regulation: The Case of GDPR

The penultimate challenge identified was related to legislation and regulation, and in particular GDPR. Some participants acknowledged some progress with legislation but argued that it does not go far enough, offering effective data protection. Legal systems and regulations are often too slow to emerge and cannot keep up with the fast pace of technology, according to some participants. It was argued that:

> 'companies are already struggling with the GDPR. If we talk about global companies, then it's even more of a struggle because, … Just like every other technology, the technology advances and then the policies start to follow.' (Focus Group A).

There was also the issue of who is going to do the monitoring and ensure that people or companies comply with the regulation (Focus Group F). This challenge was related to proper and adequate implementation of regulations and laws:

> 'On the one hand, you've got the government, and the legal perspective and the regulation, which is falling behind when we look at Facebook and how they get away with things, etc' (Focus Group A).

Some participants noted how new legislation through parliament is a lengthy and slow process (unlike the fast-paced nature of technological advances). Participants also made observations regarding power dynamics—'it all comes back to politics and power'—that ultimately meant legislation ends up protecting big companies rather than consumers:

> 'The legal systems are designed to protect the big companies, not the consumer...Yes, we have, with GDPR, these massive fines, but then all I can see that that leads to is a protracted legal battle (Focus Group A).'

Examples were given of companies that lobby legislators and specific cases such as IBM were mentioned where they just "sat it out and made things very difficult for a period of years until the case was dropped" (Focus Group A).

Limitations of legislations were also related to the argument that national and regional regulations regarding digital media do not work if they do not have a global perspective, especially given that the digital world does not have the same physical geographical borders of the offline world. Therefore, when certain countries or regions pass regulations that are legally binding, they often do not have the ability to control actions, processes and behaviours beyond their borders. A lack of global

collaboration regarding laws and regulations was perceived as a hindrance:

> 'I think it's really important to call for, sort of, a global collaboration where companies, and policymakers, and all other stakeholders involved in keeping data private, and stored in a lawful and fair way, to just make sure that this goes on in the best way possible. From what we've seen so far, this hasn't been done so well, even with the sort of conventional technologies that we've been using so far but, when it comes to AI it's even much more of a bigger issue' (Focus Group A).

Given the EU context of the study, GDPR-related challenges dominated the focus group discussions on legislative challenges. One participant argued that despite the general positive aspects of GDPR, what is lacking from it is group privacy protection that goes beyond individual data protection and looks at "how the data are being merged, are being collected, and so this kind of a connection between people ... the protection is not strong enough there" (Focus Group A). Another limitation brought forward was that it has "not even touched the surface" of issues related to data ownership, how data is sourced, maintained, managed, removed etc (Focus Group F).

The negative effect on innovation was also a tension that legislation and in particular GDPR brought about: "Things like the GDPR actually make it very difficult for companies to innovate, because of the restrictions that the GDPR puts on them" (Focus Group A). A final limitation, described as a "major" one by the participant was its inability to "cope with blockchain". The tension between GDPR and blockchain technologies relates to, for instance, the difficulty in applying legislation originally based on centralised and identifiable natural persons who control personal data, to the decentralised nature and environment of blockchain technology. It could also refer to the immutable nature of block chain transactions which may affect the rights of data subjects such as the right of rectification and erasure ('right to be forgotten') (Kaulartz et al., 2019). Again, what is confirmed is the argument that legal frameworks have not 'caught up' with the changes in technology (Focus Group A).

Health Crises Imposing Additional Considerations: Reconciling the Dilemma Between Public Good vs Individual Privacy

The last challenge was related to health crises, triggered by the Covid-19 pandemic that brought to the fore ethical dilemmas such as the one between the 'common good' and individual privacy. The majority of focus groups took place during the Covid-19 pandemic and so the topic was unintendedly yet unsurprisingly also mentioned in the majority of the focus groups (5 out of 6). Participants recognised that the pandemic presented an unprecedented ethical challenge for policy-makers and argued that it gave additional weight to address matters related to privacy. Participants emphasised the need for a 'political debate' to be had on whether it is ethically justifiable to 'give up some of our freedoms for the greater [good]'. For example, one participant implied that this may be a necessary thing to do given the current context; speed - in search for a solution to the pandemic and data contributing fast to epidemiological models—they argued, may be prioritised over getting consent from those supplying their data (Focus Group A).

Others disagreed and pointed to the fact that a physical lock down is temporary whereas the collection of data in a virtual space may be a much more long-term project; as such 'it constrains your future actions in a way that being locked down for a period of time, and then that lockdown stops, doesn't' (Focus Group A). Additionally, participants argued that data collection and tracking through smartphones in the midst of the pandemic (to be able to monitor Covid-19 cases) should concern us in terms of the individual impact this loss of privacy may have in the long-term, potentially leading to stigma and stereotyping, while others emphasised the way the algorithms helped spread fear, panic, misinformation and fake news during the pandemic (Focus Group B, E and F)).

It seemed that the pandemic exposed the sheer lack of sufficient awareness and understanding of the public on these ethical issues and as such it offered an opportunity to bring them closer to everyday debates and discourse. Therefore, although the pandemic was identified as a challenge, it was also seen as an opportunity to speed up progress on addressing ethical issues related to the digital space, as it inadvertently created a 'huge technology learning curve' and acted as a 'big wake-up call' (Focus Group A).

DISCUSSION

Our findings contribute to the academic discourse by going beyond identification of what the ethical issues are and zooming in on the more specific obstacles, tensions and dilemmas that stakeholders—such as policymakers and researchers—are faced when attempting to improve the ethical landscape of Big Data and AI, and by implication digital media. Stakeholders identified the limitations or absence of regulatory frameworks; the lack of pressure on companies; the conflicting norms and values which result in different definitions of ethics; the rise of populism and the limited critical thinking skills of both the public user and AI developers as the main challenges for addressing the ethical issues of AI and Big Data. The implications of this paper are important as progress on addressing ethical issues and protecting democracy is based on a thorough understanding of what is at stake and what is actually preventing progress in practice, which this paper contributes to. Our findings also speak to the emerging interdisciplinary research field of public understanding of science and technology (Kalampalikis et al., 2013) and provide insights to policy makers for making emerging technologies, and digital media in particular, more ethical and more democratic.

Our research has also highlighted the relationship between design and cultural and ethical norms and values. This is in line with calls for a move away from viewing technology as naturally objective due to it not being a living organism and a call for 'incorporating moral and societal values into the design processes' (Van den Hoven et al., 2015, p. 2). Technology

design should be developed 'in accordance with the moral values of users and society at large' rather than viewed as 'a technical and value-neutral task of developing artifacts that meet functional requirements formulated by clients and users' (Van den Hoven et al., 2015, p. 1). At the same time, it is important to note that there are certain limits to what designers can do in relation to the strengthening of democracy and the rule of law. As Pols and Spahn note, several factors 'are outside the control of the engineer' or 'only under limited control of engineers, such as those that lie in the realm of use and institutional contexts' and therefore it is of no surprise 'that design methods that seek to further democracy and justice tend to focus on what engineers do have control over (though not necessarily full control): the design process' (2015, p. 357).

Designers values are manifested in the products that they create, and through the use of such products, these values are then exported into society, constituting and shaping it at the same time. Given this immense power therefore, it is urgent to 'design for value' (Helbing et al., 2019). Whose values one might ask? The global heterogeneous context, with countries having different cultural and political norms was seen by our participants as potentially creating further deadlocks along the way. This can be seen as the result of an inherent and to an extent inevitable tension between the nature of the internet with its speed of information, border permeability or border defiance, and the nature of the 'real world' with border controls, national sovereignty and specific legislation within its borders. It is also a result of polarisation and of political manipulation by populist leaders and governments. The rise of populism constitutes a challenge for AI to be more ethical because populist leaders influence the public into thinking that data collection and surveillance are not a serious ethical issue. Polarisation can also be explained by the creation of 'echo chambers' as a result of algorithmic feedback loops and closed networks in social media (Shaffer, 2019; Iordanou and Kuhn, 2020). Ultimately, as part of democracy diverse voices are to be heard but some common ground needs to be reached if we are to move forward with legislation and its implementation. We suggest that further research could aim at involving even more members of the public and identifying common ground across countries and regions, as well as cultural specific challenges (see Bauer and Süerdem, 2019) that need to be addressed.

Legislation might also help to address the challenge of defining ethics and reaching a consensus, which was also mentioned as a challenge posed when addressing the ethical challenges of digital media and other emerging technologies. Participants acknowledged that GDPR is not sufficient, proposing further measures in legislation and regulation. The concerns expressed by the participants are aligned with the concerns expressed by the Members of the European Parliament, who declared that 'we need laws, not platform guidelines' (European Commission Polices, 2021). The European Commission's two new legislative initiatives, the "Digital Service Act" and the "Digital Markets Act", are steps towards creating a safer digital world by providing gatekeeping online (European Commission Policies, 2021). These new legislative initiatives aim to tackle the issue of misinformation, which constitutes according to Moghaddam (2019), the greatest threat to democracy since Second World War, because of its influence in shaping public opinion on important issues.

Another important insight that emerges from our findings is the implication of power and control. Power, as mentioned in the data, is also related to the position of the developer. Therefore, it is important to also acknowledge the varying degrees of agency conferred upon an individual by the power structures and asymmetries that characterise the working environment of companies and organisations. A case in point is the controversial firing, often unlawful and unethical, of employees who try to raise issues of ethical significance to their employers with the goal of improving the wider implications to society (see discussion about former Google employees in the Introduction).

With power comes responsibility and our participants emphasised the need for companies to exhibit responsible innovation that promotes digital well-being rather than their own vested interests. Our data voiced stakeholders concerns regarding a context which they argued is one of 'surveillance capitalism' (Zuboff, 2019), of a 'top-down culture' and manipulation. In this 'age of surveillance capitalism' as Shoshana Zuboff aptly puts it 'automated machine processes not only *know* our behavior but also *shape* our behavior at scale' (2019, p. 15). Founders of these companies knowingly and deliberately intervene 'in order to nudge, coax, tune, and herd behavior toward profitable outcomes' (2019, p. 15). Therefore, instead of the digital space being a means to 'democratization of knowledge' it often becomes a self-legitimised, discriminatory, inert and undemocratic means to satisfy the financial interests of a few powerful tech elites (2019, p. 15-16) and powerful political elites-a kind of populist surveillance or 'surveillance populism'.

Big Tech companies have been well known for having particular strategies that exploit psychological traits of human behaviour, including dopamine release, in order to maximise the time spent (and data produced) on their systems (Haynes, 2018; see also Orlowski and Rhodes, 2020). This is not to contend that technology directly causes addiction as such, but rather that technologies seem to exacerbate both the triggers and the symptoms of other, underlying disorders like attention problems, anxiety and depression (Ferguson and Ceranoglu, 2014). Ultimately, this constitutes the exploitation of human psychological weaknesses and manipulation of vulnerable emotional states at particular points in time. There are best-selling books which teach companies how to get their users 'hooked' by creating 'products people can't put down', for example the book authored by Nir Eyal in 2013, entitled 'Hooked: how to build habit-forming products'.

Manipulation becomes even more problematic when it occurs at the complete ignorance and absence of consent—and therefore at the expense—of the user. Even the provision of consent and adequate information to the consumers of technology does not absolve the producers of the moral responsibility given that often this information is not communicated in a manner that is easy to understand (for instance, when presenting terms of conditions that are tens of pages long in order to be able to access digital media). Moreover, opting out of digital services rendering people into digital hermits, although not impossible, is far less realistic in the midst of a pandemic, leading to for instance, an inability to access banking services or the exacerbation of isolation and cutting-off of the necessary social support systems of friends, family and others. In

reality, this is not even about a 'social dilemma' (see Orlowski and Rhodes, 2020), but the presence of a 'fake dilemma' of a one-way street masked as two-way. Users are effectively sometimes left with very little choice. Digital media firms have been intensely criticised for having a 'take it or leave it' approach by both academics and policy-makers alike (see Gibbs, 2018). As discussed above, we posit that a substantial share of the responsibility for technology use that reflects or is caused by manipulation of human behaviour, lies with the technology producers, both at an individual and company level.

At the same time, as our data has shown, it would be a mistake to ignore the responsibility of the user. Before using any technology or object that is part of the Internet of Things, be it a new gadget or a new automotive, we have the task, if not the responsibility, to read the manual and be aware of the risks so as to ensure the safety of ourselves as well as others. In the case of driving an automotive, it is difficult to imagine a society where the users - in this case drivers - are left to their 'own devices' to choose whether or not they will conform to the road safety rules, without any help, guidance and a common framework of regulations. Thus, one can infer that some form of regulatory control and education is required in order for users and society as a whole to function in an orderly and efficient way. Our participants were keen to emphasise that one key setback is the lack of proactive intervention; there is reaction rather than prevention.

Education, particularly media literacy and critical thinking, is a powerful tool of preventing but also of mitigating the ethical challenges of digital media and AI. It is imperative for both researchers and policy-makers alike to invest in promoting epistemic understanding and understanding of the nature of communication in digital media if we want citizens to be able to discern facts and reliable information from fake news and misinformation. Epistemic understanding supports critical evaluation of information (Iordanou et al., 2019) and consideration of multiple dimensions in a particular issue (Baytelman et al., 2020). 'If the scientifically literate citizen-consumer is important, then the epistemic questions about how credible claims make their way from a scientific community to the individuals who use those claims are equally important' (Höttecke and Allchin, 2020, p. 644).

It is possible that the sheer, and often flagrant and unashamed exploitation of human behaviour is also at the heart of why in recent years we have seen the surge of several initiatives, often stemming from higher education research labs, for instance at Stanford, Oxford or Frankfurt[8] with a particular focus on values that lead to human flourishing and digital well-being. There are also recent EU-funded projects such as SHERPA, SIENNA and PANELFIT that often collaborate together to help address the ethical, human rights and legislative issues raised by AI, Big Data and other emerging technologies.[9] There is a form of resistance, in terms of both prevention and countering, that is fighting back against the overriding tide of unethical behaviour by placing the spotlight on precisely the element whose lack of led us to the current precariousness: a human-centred ethical approach to Big Data and AI.

Another significant insight transpiring from our data is that there are always going to be certain limits to what can be done when trying to address ethical issues in Big Data and AI and prevent human rights violations. For example, education and regulation are reasonable and potentially effective tools to improve both lack of understanding and increase ethical behaviours as well as transparency, accountability, self-reflection and critical thinking - the latter in particular raising awareness about biases - but ultimately they are not a panacea. There are also technical limits that have yet to be solved; for instance, even if the desire is there to implement GDPR or 'the right to be forgotten', how do you implement this practically when one's data has been used to generate AI or machine learning and it already consists of that, it's already embedded? Deleting the data itself does not consist of the user's data being forgotten.

Policymakers and researchers alike should therefore be aware of these limitations when designing initiatives or formulating policies; unless there is deep-rooted change to the structural systems of bias inherent in a society, efforts to address the biases in Big Data and AI will remain at a superficial level. Other challenges are dependent on further technological innovations.

An interesting observation is the extent of negative language, prevalent across the focus groups that is used to describe the activities of companies, engineers, developers etc. The choice of language by itself denotes an alarming situation, a discursive gap that policy-makers may consider how to constructively bridge. The negative language towards tech companies provides information about public understanding of science. Evaluation of science constitutes one of the basic indicators of public understanding of science, along with literacy and engagement (Bauer and Süerdem, 2019). Participants in this study exhibited a negative evaluation of technology. In addition, the negative language towards tech scientists implies a lack of trust in scientists, who could invest more in their communication with the public, highlighting the alignment of scientists' values with the public interest (Oreskes, 2019) to restore public trust in science. Deliberative discussion is at the heart of democracy and this study enriches the discourse surrounding digital democracy by foregrounding the voices and perspectives of stakeholders. Further research could perhaps also be done in terms of why exactly there was this alarmist approach, delving more specifically into individual experiences but also a greater focus on what should be promoted rather than avoided i.e. how not only to prevent and mitigate the harms caused by Big Data and AI, and by implication, digital media, but also highlighting good practices that should be followed.

To better protect the foundations of democracy, we argue that there should be a better balance between direct and indirect (representative) democracy, particularly in countries which tend to operate primarily through the latter. The principle of indirect democracy and elected representatives rests on the often outdated assumption that those elected are more knowledgeable and in a better position to make decisions on

[8]Stanford University's Data Science Initiative's relevant domain is entitled 'Data Science for Humanity'. See here https://sdsi.stanford.edu/about/data-science-humanity. The Oxford Digital Ethics Lab (led by Luciano Floridi) has an emphasis on "digital innovation as a force for good". See here: https://digitalethicslab.oii.ox.ac.uk/. The Frankfurt Big Data Lab which is hosted at the Goethe University in Frankfurt has a research area on 'The Ethics of AI' and "Big Data for Social Good". See here http://www.bigdata.uni-frankfurt.de/ethics-artificial-intelligence/

[9]See here: https://www.sienna-project.eu/news/news-item/?tarContentId=822728

behalf of the public. However, there is no convincing evidence to show that members of legislative bodies are more educated in the risks of Big Data and AI or digital media than other members of the public. Therefore, alternative bodies could be created - such as ethical councils with dedicated experts from research and industry, local assemblies, as well as citizen science initiatives and referenda (for instance as part of e-democracy) - that both motivate the public to educate themselves on the one hand, and to make good use of their existing expertise on the other. If we want to create a better digital democracy, then we need to go beyond normative calls for 'making the right decisions' but create safe spaces for democratic deliberation and constructive and creative ways to accommodate and negotiate dissent and diversity.

Despite its limits, education of the public needs to happen at all ages, from early childhood to the elderly as both a strategy of prevention of harms and mitigation of risks. However, rather than merely providing the public with information about data privacy, digital literacy, the importance of transparency etc., it is important to also educate the public about alternative ways in which their data and ownership of it can be channelled, and for public services to provide the necessary tools and structures for doing so. For instance, revenues from personal health data could be collected and shared for the betterment of society rather for private companies to obtain more money than they could possibly spend in a lifetime, a crucial element that has widened the inequality gap between the rich and the poor in recent years. Suggestions have already been put forward for citizen-owned non-profit data cooperatives which puts the public back in control of their data, provides a trustworthy framework for data donation and a means for digital well-being that addresses socio-economic inequalities (Hafen, 2019; Loi et al., 2020).

Ultimately, when attempting to reach a compromise between ethics, digital well-being and democracy, the focus should not be on whether the unethical practices are a result of malevolent or benevolent behaviour; good intentions do not lessen the harm caused. What's more, engaging in scapegoating processes that merely label some as 'spoilers' are more likely to further divide than constructively change society for the better. We need to strengthen citizens' digital agency and self-determination at all levels of society, health, economic, political and social, otherwise our countries may resemble more totalitarian systems rather than democratic ones. Digital media and other technologies need to resonate with and reflect the ethics of society at large, not to be treated as if the digital world has less to do with human subjects and so erroneously presume that ethics matter less. Both far-right and Islamist violent extremism have shown that it is very easy for violent words to turn into violent actions, threatening not just internal stability within country borders but also having global ramifications, given that unlike the 'real world' the digital world has very few borders.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors upon request, with permission of the third party. The authors will need individual consent from all the participants whose data was collected in order to make their data publicly available.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Cyprus National Bioethics Committee. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KI contributed to the conception, design and data collection of the study. EC performed and reported the data analysis. Both authors contributed to writing the manuscript and revising it, and read and approved the submitted version.

FUNDING

This work was funded by a grant from the European Union's Horizon 2020 research and innovation programme (Project: SHERPA, No. 786641).

ACKNOWLEDGMENTS

The authors would like to thank all of their colleagues who participated in this project as well as the participants of the focus groups for their time and valuable input.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpos.2021.682945/full#supplementary-material

REFERENCES

ACM (2018). "World's Largest Computing Association Affirms Obligation of Computing Professionals to Use Skills for Benefit of Society". Available at: https://www.acm.org/media-center/2018/july/acm-updates-code-of-ethics.

Ali, M., Sapiezynski, P., Korolova, A., Mislove, A., and Rieke, A. (2019). *Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging*. Cornell University. https://arxiv.org/abs/1912.04255.

Amaro, S. (2021). Democracy Could Have Been Damaged Forever in the Last 4 years, EU Chief says.the Davos Agenda. Available at: https://www.cnbc.com/2021/01/26/davos-democracy-might-be-damaged-forever-eu-von-der-leyen.html.

Bauer, M. W., and Gaskell, G. (1999). Towards a Paradigm for Research on Social Representations. *J. Theor. Soc. Behav.* 29 (2), 163–186. doi:10.1111/1468-5914.00096

Bauer, M. W., and Gregory, J. (2007). From Journalism to Corporate Communication in post-war Britain, in *Journalism, Science and Society: Science Communication Between News and Public Relations.* Editors Bauer M. W., Bucchi M. (New York and London: Routledge), 33–51.

Bauer, M. W., and Süerdem, A. (2019). "Four Cultures of Science across Europe," in *The Cultural Authority of Science: Comparing across Europe, Asia, Africa and the Americas.* Ediors Bauer M. W., Süerdem A., Pansegrau P., Shukla R., (New York and London: Routledge), 301–318.

Baytelman, A., Iordanou, K., and Constantinou, C. P. (2020). Epistemic Beliefs and Prior Knowledge as Predictors of the Construction of Different Types of Arguments on Socioscientific Issues. *J. Res. Sci. Teach.* 57 (8), 1199–1227. doi:10.1002/tea.21627

BBC News (2021b). Margaret Mitchell: Google Fires AI Ethics Founder. Available at: https://www.bbc.com/news/technology-56135817.

BBC News (2021a). Twitter' permanently Suspends' Trump's Account. Available at: https://www.bbc.com/news/world-us-canada-55597840.

Boyatzis, R. E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks: Sage.

Braun, V., and Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* 3 (2), 77–101. doi:10.1191/1478088706qp063oa

Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge: MIT Press.

Bryman, A. (2008). *Social Research Methods*. 3rd edition. Oxford: Oxford University Press.

Bucher, T. (2012). Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook. *New Media Soc.* 14, 1164–1180. doi:10.1177/1461444812440159

Burr, C., and Floridi, L. (2020). "The Ethics of Digital Well-Being: A Multidisciplinary Perspective," in *Ethics of Digital Well-Being, A Multidisciplinary Approach*. Editors C. Burr and L. Floridi (Cham: Philosophical Studies Series), 1–29. doi:10.1007/978-3-030-50585-1_1

Cacciatore, M. A., Yeo, S. K., Scheufele, D. A., Xenos, M. A., Brossard, D., and Corley, E. A. (2018). Is Facebook Making Us Dumber? Exploring Social media Use as a Predictor of Political Knowledge. *Journalism Mass Commun. Q.* 95 (2), 404–424. doi:10.1177/1077699018770447

Cave, S. (2019). *To Save Us from a Kafkaesque Future, We Must Democratise AI*. The Guardian. Available at: https://www.theguardian.com/commentisfree/2019/jan/04/future-democratise-ai-artificial-intelligence-power.

Charmaz, K. (2004). "Grounded Theory," in *The Sage Encyclopedia of Social Science Research Methods*. Editors M. S. Lewis-Beck, A. Bryman, and T. F. Liao (Thousand Oaks, California: Sage).

Costa, E., and Halpern, D. (2019). *The Behavioural Science of Online Harm and Manipulation, and what to Do about it*. The Behavioural Insights Team. Available at: https://www.bi.team/publications/the-behavioural-science-of-online-harm-and-manipulation-andwhat-to-do-about-it/

Danidou, Y. (2020). "Trusted Computing Initiative on the Spectrum of EU Cyber-Security Legal Framework," in *EU Internet Law in the Digital Era*. Editors T. E. Synodinou, P. Jougleux, C. Markou, and T. Prastitou (Cham: Springer). doi:10.1007/978-3-030-25579-4_13.277-296

Diakopoulos, N. (2015). Algorithmic Accountability. *Digital Journalism* 3 (3), 398–415. doi:10.1080/21670811.2014.976411

European Commission Polices (2021). Shaping Europe's Digital Future: Policy the Digital Services Act Package. Available at: https://ec.europa.eu/digital-single-market/en/digital-services-act-package.

Ferguson, C. J., and Ceranoglu, T. A. (2014). Attention Problems and Pathological Gaming: Resolving the 'Chicken and Egg' in a Prospective Analysis. *Psychiatr. Q.* 85, 103–110. doi:10.1007/s11126-013-9276-0

Frank, R. H. (2021). *The Economic Case for Regulating Social media*. New York Times: Economic View.

Gibbs, S. (2018). *EU: Data-Harvesting Tech Firms Are 'sweatshops of Connected World*. The Guardian. Available at: https://www.theguardian.com/technology/2018/may/02/eu-tech-firms-privacy-emails-gdpr-data-protection-supervisor.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. New Haven, CT: Yale University Press.

Goyanes, M., Borah, P., and Gil de Zúñiga, H. (2021). Social Media Filtering and Democracy: Effects of Social Media News Use and Uncivil Political Discussions on Social Media Unfriending. *Computers in Human Behavior*, 120, 106759. doi:10.1016/j.chb.2021.106759

Hafen, E. (2019). "Personal Data Cooperatives – A New Data Governance Framework for Data Donations and Precision Health," in *The Ethics of Medical Data Donation. Philosophical Studies Series, Vol 137*. Editors J. Krutzinna and L. Floridi (Cham: Springer), 141–149. doi:10.1007/978-3-030-04363-6_9

Hameleers, M., Powell, T. E., Van Der Meer, T. G. L. A., and Bos, L. (2020). A Picture Paints a Thousand Lies? the Effects and Mechanisms of Multimodal Disinformation and Rebuttals Disseminated via Social media. *Polit. Commun.* 37 (2), 281–301. doi:10.1080/10584609.2019.1674979

Hartmans, A. (2021). *Bill Gates Says Trump Should Probably Be Allowed Back on Facebook, Despite His 'corrosive' Statements about the Election*. Business Insider. Available at: https://www.businessinsider.com/bill-gates-trump-facebook-account-should-probably-be-reinstated-2021-2.

Haynes, T. (2018). *Dopamine, Smartphones & You: A Battle for Your Time*. Harvard University SITN Boston. Available at: http://sitn.hms.harvard.edu/flash/2018/dopamine-smartphones-battle-time/.

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., et al. (2019). "Will Democracy Survive Big Data and Artificial Intelligence?" in *Towards Digital Enlightenment* (Cham: Springer), 73–98. doi:10.1007/978-3-319-90869-4_7

Höttecke, D., and Allchin, D. (2020). Reconceptualizing Nature-of-science Education in the Age of Social media. *Sci. Edu.* 104 (4), 641–666. doi:10.1002/sce.21575

Iordanou, K., Christodoulou, E., and Antoniou, J. (2020). *D 4.2 Evaluation Report*. De Montfort University. doi:10.21253/DMU.12917717.v2

Iordanou, K., Kendeou, P., and Zembylas, M. (2020). Examining My-Side Bias during and after reading Controversial Historical Accounts. *Metacognition Learn.* 15 (3), 319–342. doi:10.1007/s11409-020-09240-w

Iordanou, K., and Kuhn, D. (2020). Contemplating the Opposition: Does a Personal Touch Matter? *Discourse Process.* 57 (4), 343–359. doi:10.1080/0163853X.2019.1701918

Iordanou, K., Muis, K. R., and Kendeou, P. (2019). Epistemic Perspective and Online Epistemic Processing of Evidence: Developmental and Domain Differences. *J. Exp. Edu.* 87 (4), 531–551. doi:10.1080/00220973.2018.1482857

Isaak, J., and Hanna, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer* 51 (8), 56–59. doi:10.1109/mc.2018.3191268

Jangid, H., Prabhu, A., Guhathakurta, D., Jain, J., Subramanian, M., Reddy, M., et al. (2021). *Capitol (Pat) Riots: A Comparative Study of Twitter and Parler. Computer Science*. arXiv preprint, arXiv:2101.06914 [cs.CY]. Available at: https://arxiv.org/abs/2101.06914.

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

Kalampalikis, N., Bauer, M. W., and Apostolidis, T. (2013). Science, Technology and Society: the Social Representations Approach. *Revue internationale de Psychol. sociale* 26 (3), 5–9. Available at: https://www.cairn.info/revue-internationale-de-psychologie-sociale-2013-3-page-5.html

Kaulartz, M., van Kranenburg–Hanspians, K., and Sanders, S. (2019). The Tension between GDPR and the Rise of Blockchain Technologies' CMS Legal Services. Available at: https://cms.law/en/int/publication/the-tension-between-gdpr-and-the-rise-of-blockchain-technologies.

Lewis, S. C., and Westlund, O. (2015). Big Data and Journalism. *Digital journalism* 3 (3), 447–466. doi:10.1080/21670811.2014.976418

Loader, B. D., Vromen, A., and Xenos, M. A. (2016). Performing for the Young Networked Citizen? Celebrity Politics, Social Networking and the Political Engagement of Young People. *Media, Cult. Soc.* 38 (3), 400–419. doi:10.1177/0163443715608261

Loi, M., Dehaye, P.-O., and Hafen, E. (2020). Towards Rawlsian 'property-Owning Democracy' through Personal Data Platform Cooperatives. *Critical Review of International Social and Political Philosophy*, 1–19. doi:10.1080/13698230.2020.1782046

Manavis, S. (2021). *"It Has Always Been Easy for Social media Firms to Pull the Plug on Extremism"*. New Statesman. Available at: https://www.newstatesman.

com/science-tech/social-media/2021/01/it-has-always-been-easy-pull-plug-extremism-trump-twitter-ban-parler.

Moghaddam, F. M. (2019). *Threat to Democracy: The Appeal of Authoritarianism in an Age of Uncertainty*. American Psychological Association.

Müller, V. C. (2020). "Ethics of Artificial Intelligence and Robotics," in *The Stanford Encyclopedia of Philosophy (winter 2020 Edition)*. Editor E. N. Zalta. URL: https://plato.stanford.edu/archives/win2020/entries/ethics-ai/>.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

O'Connor, C., and Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *Int. J. Qual. Methods* 19, 160940691989922. doi:10.1177/1609406919899220

Oreskes, N. (2019). *Why Trust Science?* Princeton University Press.

Orlowski, J., and Rhodes, L. (2020). *The Social Dilemma*. Documentary, Los Gatos, CA: Netflix Originals.

Peters, C. (2021). Every Online Platform that Has Cracked Down on Trump. Available at: https://www.vox.com/2021/1/10/22223356/every-platform-that-banned-trump-twitter-facebook-snapchat-twitch.

Pierce, R. (2008). *Research Methods in Politics: A Practical Guide*. London: SAGE Publications.

Pols, A., and Spahn, A. (2015). "Design for the Values of Democracy and Justice," in *Handbook of Ethics, Values and Technology Design* (Springer), 335–363. doi:10.1007/978-94-007-6970-0_13

Postill, J. (2018). Populism and Social media: a Global Perspective. *Media, Cult. Soc.* 40 (5), 754–765. doi:10.1177/0163443718772186

Rodriguez, S. (2021). "Facebook, Twitter Lock Trump's Account Following Video Addressing Washington Rioters". Available at: https://www.cnbc.com/2021/01/06/twitter-pledges-action-on-any-calls-for-violence-in-capitol-riot.html.

Romm, T. (2020). *U.S. States Sue Facebook as an Illegal Monopoly, Setting Stage for Potential Breakup*. Washington Post. Available at: https://www.washingtonpost.com/context/u-s-government-and-48-state-attorneys-general-files-lawsuit-against-facebook/5b97bd6f-8d7f-4ee2-b9ea-79a1fafc7661/?itid=lk_interstitial_manual_6 (Accessed December 10, 2020).

Roose, K. (2021). *In Pulling Trump's Megaphone, Twitter Shows where Power Now Lies*. The New York Times. Retrieved from https://www.nytimes.com/2021/01/09/technology/trump-twitter-ban.html.

Ryan, M., Antoniou, J., Brooks, L., Jiya, T., Macnish, K., and Stahl, B. (2021). Research and Practice of AI Ethics: A Case Study Approach Juxtaposing Academic Discourse with Organisational Reality. *Science and Engineering Ethics*, 27(16). doi:10.1007/s11948-021-00293-x

Sammut, G., and Bauer, M. W. (2021). *The Psychology of Social Influence: Modes and Modalities of Shifting Common Sense*. Cambridge University Press.

Scheufele, D. A., and Krause, N. M. (2019). Science Audiences, Misinformation, and Fake News. *Proc. Natl. Acad. Sci. USA* 116 (16), 7662–7669. doi:10.1073/pnas.1805871115

Schroeder, R. (2018). "Digital media and the Rise of Right-wing Populism," in *Social Theory after the Internet: Media, Technology, and Globalization* (London: UCL Press), 60–81. doi:10.2307/j.ctt20krxdr.6

Shaffer, K. (2019). "Swimming Upstream," in *Data versus Democracy* (Berkeley, CA: Apress), 31–44. doi:10.1007/978-1-4842-4540-8_3

Shearer, E. (2021). *More Than Eight-In-Ten Americans Get News from Digital Devices*. Washington: Pew Research Center. Available at: https://www.pewresearch.org/.

Smolan, S. (2016). The Human Face of Big Data. *PBS Documentary* 24, 56.

Stahl, B. C., Andreou, A., Brey, P., Hatzakis, T., Kirichenko, A., Macnish, K., et al. (2021). Artificial Intelligence for Human Flourishing – beyond Principles for Machine Learning. *J. Business Res.* 124, 374–388. doi:10.1016/j.jbusres.2020.11.030

Stoller, M., and Miller, S. (2021). *"Donald Trump Being Banned from Social media Is a Dangerous Distraction"*. The Guardian. Available at: https://www.theguardian.com/commentisfree/2021/jan/11/trump-twitter-ban-capitol-attack-facebook-youtube-google.

Talmud, I., and Mesch, G. (2020). *Wired Youth: The Online Social World of Adolescence*. 2nd Edition. London and New York: Routledge.

Thiele, L. P. (2020). Politics of Technology-Specialty Grand Challenge. *Front. Polit. Sci.* 2, 2. doi:10.3389/fpos.2020.00002

Van den Hoven, J., Vermaas, P. E., and Van de Poel, I. (2015). "Design for Values: An Introduction," in *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, 1–7. doi:10.1007/978-94-007-6970-0_40

Wachter, S., Mittelstadt, B., and Floridi, L. (2017). Transparent, Explainable, and Accountable AI for Robotics. *Sci. Robot.* 2 (6), eaan6080. doi:10.1126/scirobotics.aan6080

White, D. E., Oelke, N. D., and Friesen, S. (2012). Management of a Large Qualitative Data Set: Establishing Trustworthiness of the Data. *Int. J. Qual. Methods* 11, 244–258. doi:10.1177/160940691201100305

Woolley, S. C., and Howard, P. N. (2016). Automation, Algorithms, and Politics| Political Communication, Computational Propaganda, and Autonomous Agents Introduction. *Int. J. Commun.* 10 (0), 9, 2016. Available at: https://ijoc.org/index.php/ijoc/article/view/6298

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for Human Future at the New Frontier of Power*. London: Profile Books.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.