OPINION PAPER

The AI ethicist's dilemma: fighting Big Tech by supporting Big Tech

Henrik Skaug Sætra[1] · Mark Coeckelbergh[2] · John Danaher[3]

Abstract

Assume that a researcher uncovers a major problem with how social media are currently used. What sort of challenges arise when they must subsequently decide whether or not to use social media to create awareness about this problem? This situation routinely occurs as ethicists navigate choices regarding how to effect change and potentially remedy the problems they uncover. In this article, challenges related to new technologies and what is often referred to as 'Big Tech' are emphasized. We present what we refer to as the AI ethicist's dilemma, which emerges when an AI ethicist has to consider how their own success in communicating an identified problem is associated with a high risk of decreasing the chances of successfully remedying the problem. We examine how the ethicist can resolve the dilemma and arrive at ethically sound paths of action through combining three ethical theories: virtue ethics, deontological ethics and consequentialist ethics. The article concludes that attempting to change the world of Big Tech only using the technologies and tools they provide will at times prove to be counter-productive, and that political and other more disruptive avenues of action should also be seriously considered by ethicists who want to effect long-term change. Both strategies have advantages and disadvantages, and a combination might be desirable to achieve these advantages and mitigate some of the disadvantages discussed.

Keywords Ethics · Dilemma · Big Tech · Artificial intelligence · Big Tech

1 Introduction

Assume that a researcher uncovers major problems related to the increased usage of AI and AI based social media. These problems could be related to, amongst other things, polarization [1], fake news [2], filter bubbles [3], echo chambers [4], changes in human cognition [5], discrimination and bias [6], and surveillance and challenges related to a lack of privacy [1]. When researchers discover such problems, what action should they take? This situation routinely occurs as AI ethicists navigate the choices involved in deciding how to effect change and remedy the problems they uncover. For example, if their research emphasised the grave and unfortunate consequences of Twitter and Facebook, should they promote this research by building communities on said networks? The dilemma also extends beyond social media, and entails the ethicist's relation to the major technology companies in general. If they write a book about the dangers of Amazon, for example, should they promote the book by posting links to the book on amazon.com? And should they seek work opportunities at, for example, Google or OpenAI if they are deeply concerned about the negative environmental and social implications of large-scale language models?

These examples relate to what we refer to as the ethicist's dilemma, which emerges when an ethicist has to consider how their own success in communicating an identified challenge is associated with a high risk of decreasing the chances of successfully facing the challenge. This dilemma occurs in situations in which the means to achieve one's goals are *seemingly* best achieved by supporting that which one wishes to correct and/or practicing the opposite of that which one preaches. It is thus a proper moral dilemma.

One way to approach the problem is to follow an ethical decision making process, and we structure this article after Van de Poel and Royakkers' [7] *ethical cycle*. This cycle consists of the following steps of (1) formulating the moral problem, (2) analysing the problem, (3) considering one's options, (4) ethically evaluating these options, and finally (5) reflecting and arriving at a morally acceptable action.

✉ Henrik Skaug Sætra
Henrik.satra@hiof.no

1   Østfold University College, 1757 Remmen, Norway

2   Department of Philosophy, University of Vienna, Universitätsstrasse 7 (NIG), 1180 Vienna, Austria

3   School of Law, NUI Galway, Galway, Ireland

However, deciding what constitutes a moral problem, and thus determining what decisions should go through such a process, often depends on one's personal ethical implied beliefs. To avoid anchoring the analysis in such implied beliefs, we base the analysis on a fundamental moral pluralism and the dilemma in question is examined in light of all three major ethical theories: virtue ethics, deontological ethics, and consequentialist ethics.

The dilemma is mainly relevant for the ethicists who argue, either implicitly or explicitly, for a change of the systems they examine. However, a perceived need for change in a system is not necessarily coupled with a duty for the ethicist to effect change themselves. In the rest of the article, we work from the assumption that the ethicist in question *does* feel some obligation to address the challenges they discover. The dilemma applies to the ethicist who perceives themselves so obligated, but other AI ethicists can experience the same situations as here analysed without experiencing the dilemma.

The two first steps of the ethical cycle are examined in Sect. 2, as the moral problem is identified and analysed. In Sect. 3, the examination of some of the key strategies available to the ethicist is examined, which is the third stage of the ethical cycle. In Sect. 4, the three major ethical theories are used to evaluate the options presented in the preceding section. Finally, we conclude that all three ethical theories indicate that changing the world of Big Tech using their tools or attempting to effect change through alliances with the key companies in the sector is usually not sufficient, and that it might even be morally problematic and potentially counter-productive to pursue only this strategy. Serious efforts to actually achieve change may require the ethicist to partly distance themselves from the system in question and instead, or at least also, aim at mobilising the power of politics and civil society to effect change.

2 The moral problem—the ethicist's dilemma

The problem area here discussed encompasses the negative consequences of new technologies, with a particular emphasis on artificial intelligence and Big Data. AI based social media will be a case given particular attention, but the dilemma in question also applies to other AI based technologies. The growth of these technologies has led to widespread optimism as to the potential for innovation, economic growth, and welfare, but it has also led to an increased focus on the ethically problematic aspects of such technologies [8, 9].

As there is a wide array of issues related to the phenomena here discussed, a detailed examination of them all is beyond the scope of this article. Some of the key issues

that might give rise to the dilemma relate to platform and monopoly power[10–12], social media and its effects on human relations [13], and issues related to human cognition, polarization, filter bubbles, and pressure to conform [1, 4, 5, 14]. Other issues of relevance include privacy and surveillance [15, 16], the manipulative power of tech [17, 18], algorithmic bias resulting in racist and discriminatory AI systems [6, 8, 19, 20], and the carbon footprint of AI and Big Data [21–23].

Assuming that the problems mentioned above either *are* real, or that they are at least perceived to be real by a hypothetical AI ethicist warning against them, the question becomes: What should they do about it—if anything? In the following attempt to answer this question, two assumptions are introduced to narrow the scope of the analysis and clarify what sort of situations and potential ethicists are covered by the dilemma.

The first assumption is that the AI ethicist in question has a desire to effect change, and feels a certain duty to do so, once they discover non-trivial challenges caused by AI and its applications. Issues of motivation and an understanding of the numerous roles individuals continually play thus becomes important. An ethicist is, after all, also a human, perhaps a mother, a son, perhaps a politician, and someone's partner as well. Public choice theory [24], which is used to analyse the behaviour of both politicians and voters, highlights such issues and challenges. Theorists in this school will often assume that a politician is utility maximizing, and that the variable they optimize is their own chances for re-election [25]. Let us refer to this as power-maximization. A different motivation for a politician—some would say a nobler one—might be to assume that they were mainly concerned with implementing the best policies. Let this be *cause*-optimization. Translated to the situation of the AI ethicist, we posit that the ethicist's we examine are *cause*-maximizing and not *power*-maximizing. This is not to argue that all ethicists are in fact of this type. Some may wish to increase their power within the academic community; some may care more about fame and fortune. But it is nevertheless of interest to examine the situation of someone with what might be labelled *pure ethical motivations*. If the goal of the ethicist is to gain personal power, the discovery of a problem followed by widespread use of social media to draw attention to themselves *and* the problem will be in perfect accordance with the self-serving goal of gaining both attention and potential power. However, the assumed cause will suffer, and for ethicists who are *cause*-maximizing, this will not do.

The second assumption is mentioned already in the first, and that is that the challenges in question are non-trivial. What makes the issues discussed above interesting and what raises the present dilemma is that they are considered foundational and structural, as discussed in more detail below. These issues cannot be easily fixed by making a single actor

or person correct their ways. The problems are assumed to be foundational in the sense that they are linked to deep social institutions, human psychology, or aspects of the technologies that cannot be easily remedied by small fixes. The dilemma only applies to such foundational problems, as any minor problem that is widely accepted to be a problem and easily correctible, can be fixed without having to solve the moral dilemma which involves deciding on how to relate to Big Tech and their infrastructure.

It should be noted that the issue at hand is not unique to the domain of AI and Big Tech. For example, problems of 'green-washing' have led to recurring debates about the role of environmental ethicists [26]. An environmental ethicist concerned with climate change could, for example, be tempted to take a job as an ethics officer in a major oil company, arguing that they will as a result have better access to executives, the resources of oil companies, etc. and hence be in a better position to effect change. Similarly, a proponent of animal rights could consider employment at a zoo, as this will allow them to work directly on improving the conditions of specific animals in captivity. However, as will be argued in Sect. 4, there are a number of dangers associated with choosing to *borrow power from the sources of the problem*.

While related to dilemmas experienced by ethicists in other fields, we argue that the dilemma is particularly salient for those who work in the field of AI ethics. The dangers of AI are often connected to what is called 'Big Tech', usually referring to the big four or five tech companies [27–29]. The four main companies are sometimes referred to as *GAFA* (Google, Amazon, Facebook, and Apple) [30], while *GAFAM* or *MAGNAF* includes Microsoft and Netflix [29]. However, unlike, for example, big tobacco, whose power is mainly linked to its access to resources and networks not directly relevant for communicating the challenges created by the tobacco industry, Big Tech has both resources *and* control of the tools and digital infrastructure required for communicating and raising awareness about the problems that concern the AI ethicist. In this sense, the AI ethicist is more akin to the environmental ethicist concerned with the climate, and offered a job in Big Oil and a private company car and jet. Digital infrastructure and social media, for example, can be likened to cars, jets, and the fossil fuels on which they used to run. Using such tools enables the ethicist to travel far and wide with their concerns, but doing could also exacerbates the problem with which they're concerned. The AI ethicist must similarly consider not just whether or not to accept employment or resources from Big Tech, but also whether or not to use the infrastructure and tools they control for their own purpose.

Finally, the *system* we speak of in the following refers not just to individual AI companies, but to the larger eco-system of companies, social structures, and political arrangements that generate the negative impacts in question. Surveillance capitalism is a prime example of such a system, based on collected and monetising/actioning personal data. It is enabled not just by individual companies but by the economic, regulatory, political and social system, and we argue that you cannot correct surveillance capitalism simply by eliminating, for example, Facebook. You have to change the entire business model and the supporting legal–political infrastructure. While some might argue that the problems related to AI are the problems of isolated individuals and contexts, we assume that the issues are structural, and that efforts to mitigate negative impacts require working to change a system, and not just to correct or remove individual actors.

To sum up and simplify the moral problem:

AI, and AI based social media in particular, negatively impacts individuals and society in non-trivial ways, and the AI ethicist feels obligated take action to counter these impacts.

3 The ethicist's two main strategies

The next step in the ethical cycle is to consider the avenues of action available to the concerned AI ethicist. For clarity and to facilitate a straight-forward analysis, we reduce the strategies available to two: (a) working from within the system and (b) working from without. In brief, working within the system refers to both working for the tech companies that are integral to the system and/or to actively use the technologies on which the system is built. Effecting change from the outside entails abstaining from using these technologies, and choosing to conduct research and seek change either independently or by working for actors who are not part of or closely linked to the system described in the end of the previous section.

3.1 Changing the system from within

The first strategy entails somehow working *with* that which one wants to change. The reasons for this could be varied, but one key attraction of this strategy is that it can be both lucrative and provide the ethicist with power and reach. When the reason for working with those who constitute the system is to change the system, however, great care must be taken. One approach to this strategy is *parasitism* [31, 32], and another is to seek *constructive complicity* [33, 34].

First, the AI ethicist could aim at becoming a sort of 'parasite' that utilizes and takes what they can from the system while ideally not giving anything back to the system, if this entails strengthening the system [31, 32]. More insidious versions of parasitism could also entail actively attempting to subvert the system, and not merely trying to limit one's own contribution to it. This article stops short of evaluating

strategies that cross over into sabotage, however, and both working from within and without will be assumed to entail adhering to laws, and agreed-upon codes of conducts, user agreements, etc.

Second, those who enter the system to change it for the better run the risk of complicity. One example is how the French Algerian liberals during the Algerian war (1954–1962) were portrayed as *complicit* with colonial rule [35]. Another is the argument that feminists and women working in NGOs, for example, draw on, support, and in some way strengthen western and middle-class patriarchal values [33]. While complicity is a major danger when seeking change by aligning with power, those aware of this problem seek *constructive* complicity to reflexively and responsibly narrate this difficult landscape [33, 34].

Microsoft, for example, has a number of ethicists associated with them, and the same applies for all the large tech companies. While some, such as Meredith Whittaker, have left tech to continue their research in academia, many others are still chipping away at the system from within the corridors of power [36]. Timnit Gebru tried the same with Google, but after authoring an article on the dangers of large-scale language models [21], she was seemingly forced out of her job [37], providing what appears to be a clear example of the limitations of ethics-from-within. Furthermore, a large number of AI ethicists are active users of social networks controlled by companies they criticise. Twitter, for example, is a particularly popular platform for AI ethicists.[1]

A potential problem with working from within is that radical change can be hard to achieve by working from within the system. As Audre Lorde famously stated, 'The master's tools will never dismantle the master's house' [38]. A system consists of a set of institutions and an embedded logic that inevitably shapes any possible move towards change. Arne Næss, for example, argued that while deep ecologists must learn the language of economists to be able to both understand and debate them, they must be wary of becoming trapped in the logic inherent in such a language [39]. But this is hard. Systems shape perceptions both of what is possible and desirable, and any change effected from within will thus partly be a result of the logic of the system in question. Working from within the system, ethicists will arguably be prone to chase relatively minor problems and to become preoccupied with the technical details of existing solutions. Working with and within the system might make

it might make it difficult to perceive and understand the *real* problems.

Summing up:

Strategy 1: Seeking change by allying with Big Tech or actively using their technology and infrastructure.

3.2 Demolish the old and make way for the new

Some problems require changes of the very foundations of the system in question, and not simple tweaks. Just as Arne Næss [39] argued that solving the ecological problems that concerned him required a change of 'basic economic, technological, and ideological structures' [39], the same could be said to apply to the problems associated with AI and Big Tech.

This route towards change requires that an ethicist steps outside the system to properly assess and observe how it can be undermined, and which alternative strategies exist. By adopting such a strategy, the path of change is no longer necessarily determined by the logic of the existing system, and the ethicist is free to consider radically new and different societies and structures of societal and technological interaction.

As such a perspective is adopted, new avenues of achieving change are found in the social and political domains—for example, in the world of education, law and regulation. From outside of the system, fundamental evaluations of social values may provide us with a chance to consider and apply both top-down and bottom-up mechanisms for solving the problems at hand. Næss particularly emphasised the need to inform the general public and to instil in everyone a sense of duty to effect positive change. Simultaneously, however, top-down approaches involving politics and regulation might be required. These considerations apply whenever it might be beneficial to approach the current system from the outside to both understand the nature of the system and to effect change.

However, a potential drawback of this avenue is that the goal of achieving social change is not reached, because the degree of interaction with the system and the rest of society is low, so that no or only a weak alliance can be forged: the innovation is in this case limited to a marginal or local part of society (for example a small, marginal community), without in the end influencing the system very much. This is the danger of *marginalization*, *impotence* or *ineffectiveness*.

While a low degree of interaction can hinder efforts to effect change, we argue that the current very *high* degree of interaction between Big Tech and the political and legal institutions is even more problematic. The tight linkages between the system and the political domain, which we argue is crucial for controlling and correcting the system, when necessary, makes it very hard to find space 'outside'

---

[1] Further analysis on the dilemma will also require that one distinguishes between the different social networks and the effects associated with social networks in general and those associated with a particular network. Furthermore, in the later stages of the analysis, one must also consider how the use of one network impacts the potential strength of particular other networks and social networks in general.

the system and work to reform it at a fundamental level. We argue that this is one of the major issues with AI ethics at the moment. Therefore, much of the Big Tech economy is dependent on the system, and so much of society is dependent on it, that it is hard for the AI ethicist to find sufficient support and space outside the system.

To avoid this, one needs to be very strategic in terms of building alliances, while at the same time avoiding going too much into the system (with all the disadvantages that follow). This is a very fine balance to strike, and shows that in practice the two strategies might be combined or mixed to some extent. We will return to this point.

Summing up:

Strategy 2: Seeking change through distancing oneself from and marginalizing the sources of the problems and allying with other sources of power to effect change.

4 Evaluating the ethicist's options

The purpose of an ethical cycle is to arrive at an answer to the question of what someone *should* do (i.e., which of the two strategies outlined above should they adopt). However, saying that someone should do something implies that a normative evaluation has been performed and that those who say that it should be done make a statement concerning a duty to perform a particular action. One maxim—or rule by which to act—might, for example, be 'One should practice what one preaches', which would entail not using social media if the ethicist's message is that social media's negative implications are significant and outweigh its benefits. However, there are countless such potential maxims to act by, and different ways to evaluate both the maxims and actions themselves.

Answering what the ethicist should do, without implicitly smuggling in our own ethical inclinations and philosophical underpinnings, thus necessitates the explicit use of the mainstream ethical theories most often used to guide and evaluate ethical behaviour: virtue ethics, consequentialist ethics, and deontological ethics. Seeing the dilemma in light of these three ethical theories helps us disentangle some of the considerations involved in deciding how to act when facing the dilemma.

Our approach is, consequently, based on a moral pluralism premised on the idea that insight from analyses based on all three theories are required to understand the ethics of any action. Unlike Vallor [40], for example, who argues that the problem associated with deontology and consequentialism are so great that one must turn to virtue ethics, we find each of the theories to have both merits and flaws and

initially consider them complementary rather than mutually exclusive.

4.1 Virtue ethical considerations

Virtue ethics involves focusing on what characterizes the ethical person. Rather than focusing on the consequences of actions, or a set of duties or rules, the virtue ethicist emphasises the virtues associated with a moral character, as seen in relation to what is conducive to ethos. Rather than focusing directly on what to do, the virtue ethicist considers what sort of person one should be [41]. Certain character traits, such as courage, honesty, and benevolence are considered *virtues* conducive to moral flourishing [42]. Vallor [40] uses virtue ethics to examine what she refers to as *technomoral virtues*, which are the virtues required to live good lives in our era of technical mediation.

From a virtue ethics standpoint, a crucial consideration for the AI ethicist relates to the problem of not practicing what one preaches. Doing so might both (a) be wrong in itself as it is not conducive to a good life, and (b) decrease the chances of successfully effecting change, because one is perceived to be hypocritical (a moral vice; not a virtue). The latter points towards the consequences of not being virtuous, and can thus be argued to belong to the domain of consequentialism, discussed below. However, all virtue ethical evaluations have a tinge of consequentialism embedded in them, as one is concerned with discovering what results in a good life.

Moving to the context of AI and social media, assume that someone argues that privacy is a public good and that one person's lax privacy protection inflicts harm on others [43]. This person arguably casts themself in a strange light if they disregard their own privacy. For example, if this person continues to use Twitter to promote their critique, to engage in discussions, and to network with others, they are partly responsible for encouraging others to expose themselves, and also for more directly facilitating the collecting of relational data through their own activity. If one uses social media to talk with and about one's friends, data about oneself and one's friend can be gathered. Such behaviour would signal either (a) the willingness to inflict harm on others, and/or (b) that the ethicist does not actually believe what they argue to be true. Either option entails seriously undermining the ethicist's ethical appeal—*ethos*—and legitimacy [44].

While one might feel that rationality and logical arguments should be what determines the effectiveness of a message, rhetoricians have long understood that considerations about the sender of a message also matters. What is referred to as ethos—which is separated from logos (appeals to reason) and pathos (appeals to emotion)—is often divided into three aspects: sound sense, moral character, and benevolence

[44]. A reputation for sound sense can partially be established through credentials and a history of demonstrating that one has knowledge of particular topics, and the perception of an ethicist's moral character might also be negatively affected by the ethicist's actions. In this context it is worth considering both ethos as pre-established through reputation and ethos established through ongoing actions and communication. For example, if the ethicist writes books about the dangers of social networks and privacy harms, yet actively use these networks and disregard their own advice, this raises questions both about whether the ethicist actually mean what they write and why they do not practice what they preach. Scholars, such as Carissa Veliz, for example, are very critical of surveillance capitalist technology [15], yet still use social media very actively to promote their work. They could argue that they are only sharing *professional* information online, and that they are consequently not exposing their own personal data. Nevertheless, there are problems associated with actively using and implicitly encouraging others to use, the very technologies they seem to argue that we should undermine rather than support. Ethicists become role models of sorts, and might even draw new users to platforms such as Twitter by their very presence and activity.

If we, for example, accept that courage and honesty are virtues, this might imply that the virtuous person will communicate their criticism of Big Tech despite, for example, fear of retribution from the same companies. However, identifying which course of action results in following a virtue like courage is an arduous task, and associated with much practical difficulty. Nevertheless, certain virtues, such as courage, integrity, and honesty, all seem to point relatively clearly towards acting in ways that maximize the chances of effecting change and rectifying problems in the long-term, even if such action are associated with short-term and/or individual costs. But this alone is not enough to solve the dilemma, because merely acting in accordance with virtue, without an eye to how positive change and good lives *in general* will be promoted, is relatively aimless. This leads to considerations of which consequences accompany our choices.

4.2 Consequentialist considerations

As the name implies, *consequentialist ethics* is focused on the consequences of our actions. According to a popular version of consequentialism—utilitarianism—what should guide our actions is an evaluation of what will create the greatest amount of happiness for most people [42]. Often associated with the idiom that 'the end justifies the means', a utilitarian might, for example, justify using Twitter to condemn Twitter, if this will in fact achieve a general increase of happiness relative to abstaining from using Twitter for

the same purpose. The utilitarian can live with the dilemma here discussed if the means (use of the system) provided the chance to achieve the ends (change of the system). However, they will also have to grapple with the possibility that using the system will, as argued below, end up strengthening the system. The overall effect of attempting to effect change from within, then, might be counterproductive in terms of achieving a general increase of happiness, since a strengthened system will exacerbate the potential harms. Vallor [40] dismisses consequentialism (or utilitarianism), because she argues that technologies open so many new alternative futures that the consequences become 'simply incalculable'. However, consequentialism need not be premised on the need to achieve full and certain knowledge of all consequences. Rather, it could be a pragmatic approach that simply argues that we evaluate the morality of an action based on the *known*—or knowable—consequences.

A key reason for using social media (or the system more generally) to change the system and alleviate purported harms is that the system and its various tools are highly effective. While it might be commendable to practice what one preaches, it might also be ineffective, as an AI ethicist that absents themselves from social media and related technologies would, arguably, remain unseen and achieve nothing.

Being able to effect outcomes requires *power* of some sort, and the allure of Big Tech is mainly based on its power. Power comes in various forms, and both power-*over* and power-*to-affect outcomes* are relevant for this analysis [45]. As we will see, Big Tech has the power to directly impact individual AI ethicists, and this gives rise to fears of negative consequences and sanctions, etc., but also to considerations of the gains to be had from cooperating with these companies. The large technology companies are in a position to make or break many an effort to get a message to the public. Just like a government critic in a state without any protection against the arbitrary abuse of power, the ethicist might perceive themselves to be at the mercy of Big Tech. For example, a book that is 'buried' by Amazon is not very likely to make much headway in terms of sales [30]. Such considerations fuel the neo-structuralist arguments against Big Tech companies, based on considerations of Big Tech's structural monopoly-like power which create potential individual and political harms [29].

4.2.1 Power and consequences

Power is a key concept for understanding the means to effect change. It allows us to examine the power of Big Tech, but also to better understand how the ethicist might amass and exercise power to change the system. For example, the companies control resources and infrastructure that enables the ethicist to effect outcomes, which is the very reason the

dilemma emerges. Power comes in different guises, however, and can be split into *episodic, dispositional, structural,* and *constitutive* types [46]. The first three are particularly interesting for evaluating the consequences of actions, while constitutive power in this context is mainly relevant in relation to the generation of harms illuminated by the AI ethicist. Allow us to explain.

Episodic power is power someone has to influence someone else in a particular situation. It is a clear example of power-over another, and it is both context-dependent and relational. For example, to what degree can the executive at, for example, Apple use their power to silence or remove an employed ethicist that is trying to raise awareness of problematic issues in the company? Such relationships highlight the importance of power differentials, and the importance of focusing on relative, and not absolute, power. Episodic power is often contrasted with dispositional power, which is the power someone has to effect outcomes regardless of others, and even disregarding context to a certain degree. It is relational in the sense that effecting outcomes tend to involve making things change in the world, but it is not specifically aimed at relations between humans.

Both these forms of power are relevant for the AI ethicist, and if one simplifies, the companies' episodic power is of interest in that it can both hurt and aid the individual ethicist in their efforts to spread their message, while their dispositional power relates more closely to their control of resources and infrastructures as means to achieve ends.

In this sense, *systemic* power also becomes relevant, as this relates to the distribution of power and ability, or ableness [46]. One of the challenges with the system discussed here is that Big Tech holds much power, while others are perceived to be relatively powerless, or exposed, to this power. This relates both to the precarious situation of ethicists employed by or otherwise under the power of those they seek to scrutinize, and it also relates more directly to how Big Tech has power and generates the negative individual and societal outcomes that the ethicist seeks to address.

Finally, constitutive power relates to power to construct, shape, and constitute our *selves*. Focault is a philosopher often mentioned as an analyst of this kind of power [46], and in the context of this article, this sort of power is mainly interesting in that the constitutive power of Big Tech is one of the key reasons for many of the challenges discussed in Sect. 2.

The various powers of Big Tech are attractive to anyone who desires to change the world for the better, and this entices many to attempt to harness this power for good by working within the system. In the following three sub-sections, the key dangers related to the consequences of 'borrowing power' from this system are elucidated. They are all related to the problem of *complicity*, and the problem of in some way being dependent upon or part of that which one criticises. The consequences of not acting in accordance with the virtues, and consequently undermining oneself, has already been discussed in the previous section, and will not be repeated here, but this must also be factored in when consequences are used to evaluate actions.

4.2.2 Supporting the companies and making their products more attractive

Whenever an ethicist uses a social network to garner support for their message—or themselves—they are supporting the social networks in question. Even if the network is used exclusively for the purpose of discussing the posited flaws of the network, the activity generated makes the network more attractive. The more people involved, the more attractive it becomes. The ethicist who intends to change and undermine the system will using the system potentially end up strengthening it. More people will be attracted to the network, and the people involved in the debates generated will spend more time there. The ethicist and their followers could *increase* the value of the product.

In terms of consequences, then, the potential for making people aware of challenges must be evaluated against the negative consequences of simultaneously strengthening the system that generates the problems. Whenever these consequences are substantial, this suggests the ethicist should not use Big Tech infrastructure, despite its seductive appearance.

Of some importance here is the notion that the ethicist will have to consider the trade-offs between short- and long-term gains. While short-term gains might be had by spreading the message on a social network, the long-term consequences are the strengthening of the network and an exacerbation of the problem.

4.2.3 Directly supporting the platforms by driving sales of books, ads, etc.

A danger closely related to the one just discussed is that attempts to harness the tech for good involves strengthening the financial situation of the companies behind them rather directly. When an ethicist drives traffic on the social network, the owner of the network increases both their direct ad revenues and the amount of data they can gather, utilise, and sell. When the ethicist writes a book about the dangers of Amazon and promotes the link to the book on Amazon.com, Amazon will make more money, gather more data from customers the ethicist incidentally herds to their site, and thus helps to fortify their position in the market.

Once again, the consequences must be evaluated on the basis of the ethicist's goals. If the goal had been individual gain, using these platforms would naturally be legitimate. But if the *cause* is perceived as most important,

abstaining—and taking a potential personal loss—cold be the most ethically sound option.

4.2.4 Legitimising the platforms and the problem of complicity

In addition to these rather direct forms of support, the ethicist who borrows questionable power to do good, simultaneously makes the power appear to be less questionable. Using Facebook, Twitter, and Amazon to criticise these very sites and actors, they are in fact legitimising them. This mainly occurs through showing that (a) the ethicist does not consider the platform or company too problematic for them to use it/them, and (b) by directly proving that the companies are open to criticism. The former is the problem of complicity discussed above, and while some argue that constructive complicity is possibly [33, 34], this is a difficult landscape to navigate, and this will always be associated with high risks of being used as an ethics alibi in a form of 'ethics washing'.

The latter is, arguably more interesting and highlights a perverse consequence of using a tool to criticise a tool. The ethicist who critiques Big Tech and/or their tools, but who joins Big Tech or use their tools to do so, demonstrates that the system allows for criticism. This gives the companies an appearance of open-mindedness and ethical awareness. If the companies behind these sites are willing to use their immense power to support criticism of them, then surely, they cannot be so bad? Thus, when ethicists provide the companies with such a valuable service, the very criticism they attempt to communicate is weakened.

4.3 Deontology and universal rules

Deontology is often associated with Kant and the idea that we should check whether the actions we consider could be turned into universal maxims that guide the actions of all—not just ourselves in a particular situation [42]. A deontologically inclined ethicist who has just written a book about the moral rules we ought to follow to challenge the power of Big Tech, for example, should in theory abide by the same rules themselves, unless some universalizable version of the rules in question justify other avenues of action. Vallor [40] dismisses this theory, because she argues that new technologies will be ubiquitous, and it thus becomes impossible, she writes, to use rule-based ethics to evaluate actions. However, even if we might all live with a pervasive lack of privacy, many distinct ways to deal with this will be available. Furthermore, many different maxims for either working with Big Tech or refusing to do so will exist, thus making deontology a clearly viable ethical theory also in the age of advanced machines.

Consequentialism is intuitively attractive, since it provides justification for just about any action that can be assumed to

lead to one's accepted goal. One challenge with this approach is that the ethicist can justify defying most intuitive norms and maxims, since the end justifies the means. The question, then, is: what happens if everyone acts according to such a logic and make exceptions for themselves? We might then easily find ourselves in various tragedies of the commons situations [47], where too many act in ways not in accordance with what they argue to be necessary to fix a problem. Furthermore, the ethicist turns themself into a bad example and establishes precedence for breaking the advocated norms. Just like Al Gore was criticised for flying private jets to spread the message about climate change [48], the AI ethicist using Facebook to spread the message about the dangers of social media is in a vulnerable position, as their actions seem to break with any intuitive maxim, derivable from their message.

Jarod Lanier [49], for example, published the book *Ten Arguments For Deleting Your Social Media Accounts Right Now.* He serves as an example of practicing what one preaches, as he does not have (known) social media accounts. Without making assumptions about his ethical inclinations, this serves as an illustration of how an author might be argued to have proposed a maxim that it would be somewhat paradoxical if they themself did not adhere to. Some might be tempted to construct rules such as 'Don't use social media unless the consequences of using social media outweigh the negative effects', but this is actually a consequentialist argument. A deontologist, or at least a Kantian, will have to adhere to universal rules, in which some fundamental rules restrict our ability to make certain trade-offs justified by appealing overall consequences. Of particular importance is the idea that absolute moral duties—not using other people as means being the most famous—preclude certain solutions in which harms that befall some individuals are argued to be justifiable because of the benefits that befall others. One problem with this approach is that the common good will at times have to be sacrificed on behalf on some abstract notion of what constitutes just action [41].

The important point here is that the AI ethicist cannot, according to this approach, say that the end justifies the means, and use social media to warn against social media. If this was made a universal rule, *all* would potentially use social media in such ways, and no progress towards any goal would be made. More specifically, using social media to generate awareness through one's social network might also be construed as a way of using those in one's network as a means to the end for addressing a larger problem. If the social networks in themselves are harmful, such use of others cannot easily be turned into a universal principle for action.

5 What should the ethicist do?

While the theories mentioned above are often described as contrasting, Nolt [41] argues that to 'decide well' we have to think both deontologically and consequentially. We have added virtue ethics into the mix as well. This entails of form of *moral pluralism*, which, in brief, is argued to be the only rational response to the fact that no ethical theory alone encompasses all possible contexts, subjects, and processes [50]. This pluralistic approach is in line with pragmatism, in particular with Dewey's ethics [51], which refuses to reduce moral reasoning, experience, and wisdom to a single principle or theoretical source.

This section addresses the question of how the ethicist can use the ethical theories to arrive at a choice between the two available strategies. The merits of each strategy are briefly evaluated, before we show how the three theories tend to converge on a recommendation that both strategies must to some degree be followed simultaneously while emphasizing the need to focus sufficiently on strategy 2. Finally, the nature and difficulty involved in the dilemma will change depending on both the context and the situation of the particular ethicist who faces it, and the need to also consider actual people is addressed in Sect. 5.3.

5.1 The potential and limitations of strategy 1

The most obvious benefit of strategy 1 is that, when successful, it provides a way to be close to the sources of power and to guide and change the exertion of such power in a beneficial way. When this can be done, it is both less conflictual and seemingly more effective, as the closeness to those with the power to make change happen makes direct dialogue and persuasion possible.

The power issues can also be described by means of what one may call a relational approach to power, perhaps in combination with the constitutive approach to power advanced by Foucault. Power is not necessarily something bad or limiting. It can be enabling. And it is deeply social. An individual's power always depends on the power of others, which may limit us but may also enable us. We already talked about making alliances. The success of either strategy will depend on how the AI ethicist's power interacts with the power of others. In the first strategy, the ethicist needs to ally with powerful people within the system to reach their goals. These powerful people may limit the ethicist's power, but in the best case they can support the ethicist. In the second strategy, the alternative proposed by the ethicist can only find wider implementation in society if the ethicist allies with powerful actors outside the system. Individuals alone cannot change the system; given the relational nature of power, alliances need to be made—inside or outside the system. Furthermore, technologies can enhance one's powers, including the power to effectuate social change. They should not only be seen as 'the instrument of the enemy'.

This approach differs from the allergy proponents of the second strategy may have for 'power' or 'powerful' actors. Power is not dirty; in both strategies, it is needed for enabling social change. It only becomes 'dirty' when the exertion of it is illegitimate [1]. This is why we discussed ethical theories and their pragmatic uses in this context.

The limitations of strategy 1 has been thoroughly covered, and they include a potential for exacerbating the problems sought addressed and for blocking other means of achieving change through legitimizing the actors in the system and marginalizing ethicists pursuing strategy 2. Furthermore, if the problems are significant enough, allying with those causing the problems can reflect poorly on the ethicist and make them appear to be both hypocritical and self-serving rather than virtuous agents of positive change.

5.2 The potential and limitations of strategy 2

All three ethical theories, in combination, provide good reasons for seriously considering strategy 2 and working to effect change from without the system.

One reason is that *ethos*—the perceived virtue of the ethicist—is important for being able to achieve desirable consequences. Stating this is not the same as stating that the morality of an author is relevant for the evaluation of the validity of their claims. Drawing on Merton, one might argue that knowledge and value are clearly separated, and that the morality of person *X* has no impact on the validity of the truth claims *X* makes [52]. Nevertheless, the effectiveness of the ethicist who does not practice what they preach will be reduced, as their ethos suffers, hurting their ability to effectively communicate their research and motivate people to effect change [53].

But strategy 2 can be argued to be unattractive due to its difficulties and the need for a long-term mindset. The ethicist might be perceived as successful if they attract much attention and support in their efforts to decry Big Tech through the use of Big Tech tools, but while their *individual* happiness might thus increase, this will come at the cost of the harms produced by their actions as the system is strengthened. Strategy 1 could thus be argued to represent a form of selfish short-termism. This dilemma is made even more difficult by the fact that the private benefits associated with for example social networks are quite tangible, while the negative collective consequences are harder to perceive. This has led, for example, to the proposal that privacy is a public good prone to market failure when individuals are free to act on the basis of their personal evaluations of the consequences for themselves alone [43].

Specifically, a tech company could hire prominent and upcoming AI ethicists in an effort to achieve several goals at the same time. First, AI ethicists provide a certain degree of legitimacy and seemingly demonstrate that the company takes ethical concerns seriously. 'Ethics washing' (or *green-washing* in terms of environmental ethics) is a term used to describe how companies will often superficially adhere to ethical principles to avoid a blowback from the public, or, worse, strong regulation. By allowing themselves to be used in such a way, ethicists may even invite the undermining of the field of ethics itself [26, 54].

Furthermore, the AI ethicist becomes entangled in an incentive structure that makes efforts to achieve radical change both risky and often self-defeating. While being employed by Big Tech does not invalidate one's research, it seems likely that solutions proposed by those in such positions are more focused on changing and adjusting the current system rather than overthrowing it.

While the common good might be sacrificed by a deontologist, a common objection to utilitarianism is that it can legitimise various forms of injustice if it is done to achieve a common good [41].

Rather than working within the logic of the existing system, we could heed the advice of Næss, who argued that we must exercise our right to evaluate technological progress and fundamental values through our political systems [55]. He argued that even if our democratic systems have abdicated much power to multinational corporations, hope is not lost, and the power is there to be taken back.

The system itself functions according to a set logic, and while progress seen from within the system is determined by whether it contributes to the effectiveness of the system, we might need to step outside it to determine whether or not the system itself is really good for humans, our societies, and our environment [39].

The political domain is where foundational questions about what sort of society we desire are answered, and it is also the source of the power required to change the system [56]. By regulation and law, the systems in place can be effectively changed. The structures that allow Big Tech to gather data as they do now are contingent; they are not essential or necessary to the social and economic environment. For example, there is no necessity associated with companies being allowed to operate in legal grey areas both nationally and internationally [57].

It is crucial that we recognise and learn from how the political domain has been used to handle massive problems before. We regulate telephone operators, for example, and their use of private data. We also regulate the food industry, transportation, etc. Come to think of it, there is hardly *any* industry that is not heavily regulated and shaped by politics. Granted, there will always be costs associated with regulation, and dangers involved in providing the

government with power. The dangers of government power are real, but, as noted above, power is not always a bad thing and the dangers associated with unbridled innovation and application of new technologies must be considered [1]. These are all reasons to favour strategy 2.

5.3 Individual contexts and considerations

While the preceding considerations have painted a relatively unflattering picture of anyone choosing strategy 1, it is also necessary to consider certain factors that might help explain why many might be tempted to choose this strategy. This is where individual's situations and contextual factors enter the equation, and where the systemic power of Big Tech becomes obvious.

What makes the dilemma particularly acute for some is that, for example, AI ethicists often have their own legitimate personal goals and ambitions to consider in addition to their ideological goals and the ideas proposed in their research. For example, an early career researcher needs publications, and they'll need their publications to be noticed, and preferably cited. Such is the road to the promised land called tenure. However*,* whether their ambition is to succeed in academia or in the private sector matters little: the need to get noticed and to connect with others who can provide future opportunities are strong, and both are hard to achieve without the tools provided by Big Tech. Twitter and LinkedIn (Microsoft) are examples of what many consider crucial tools for these tasks.

In addition to the need to get noticed and to network, there is the aforementioned danger involved in antagonising those with power. If one criticises Microsoft too strongly, there is a chance that the next job opening at the company slips a bit further away. Many ethicists will be attracted to such positions. Tarleton Gillespie, for example, is a prominent scholar of platform power, now employed by Microsoft [11]. Danah Boyd and Kate Crawford have written critical articles about Big Data, while also being employed at Microsoft [58]. Timnit Gebru, a pioneer in AI ethics who recently co-authored an article on the danger of large-scale language models [21], was until recently employed at Google. Abebe Birhane has, for example, revealed bias in widely used datasets [59], and was recently employed by DeepMind. These examples serve to illustrate that the challenges discussed in this article are real and actually apply to prominent ethicists every day.

Furthermore, those not in particularly privileged positions by necessity care about their monthly income, and antagonising the most powerful players in the industry is consequently quite unattractive for aspiring ethicists, as their jobs pay quite well. Getting invited as a speaker to the next big event sponsored by, for example, Microsoft might become less likely if you pursue strategy 2. This

need not even be the result of direct action taken by the company, but could simply happen through, for example, the organisers having a desire not to potentially anger future or current sponsors. These are the chilling or silencing effects previously mentioned.

In short, there is a dilemma involved in balancing whatever ideological commitment one has regarding the dangers of Big Tech and the personal ambition for a career either in academia or in the private sector.

6 Conclusion

Who today dares to deny the academic's incredible and almost uncanny responsibility for our society's future development [55]

The current sociotechnical system, described by Shoshanna Zuboff [57] as *surveillance capitalism,* bears a striking resemblance to the industrial society described decades earlier by Herbert Marcuse [60]. In such a system, a comfortable unfreedom prevails, and the system is exceedingly good at neutralising any efforts to disrupt it by making such efforts less effective by turning them into something compatible with the system. Just as Marxists may lament the emergence of social democracy, as it potentially pacifies the working class and prevents their revolution from unfolding, so could the AI ethicist's efforts to effect change from within the system be seen as an activity that undermines the possibility for real change. It arguably also makes it ever harder for those who attempt to effect radical change from outside of the system.

In the preceding pages, a number of challenges related to the work of AI ethicists have been examined, with a particular focus on highlighting some of the fundamental problems involved in attempting to participate in solving the problems and issues one identifies and uncover through AI ethics research.

The ethicist's dilemma arises as soon as the desire to effect change is seemingly most easily satisfied using the very systems that needs changing. In this article it is shown that the dilemma involves either strengthening the system by attempting to harness its powers, or potentially not achieving anything by relinquishing the means of using technology to spread one's message. An environmental ethicist who is sincerely concerned about the effects of climate change could start working as an ethics officer for Big Oil, but there is a chance that doing so may 'trap' them in both a logic and an incentive structure that make real change hard to achieve. An AI ethicist contemplating the dangers of new technologies is faced with a similar problem, when they are, for example, offered a lucrative job at a Big Tech company

with a quite uncertain future outside the mainstream as the only alternative.

Turning to the practicalities of change, some rightly argue that political power is dangerous [61]. Furthermore, they might argue that private initiative and innovation is the key to the good life and human welfare. However, the dangers of technology and unbridled innovation are also real. At least according to the ethicists. And if they are serious about these dangers, it may be necessary to emphasise the political domain and its power to disrupt the technological system. The dangers of private power must be bridled by the power of government, and this is in a sense a liberal argument in favour of more active use of government power [1]. Private companies generate a range of problems, and when these are understood as problems resulting from a too free market, government intervention for the sake of correcting market failure is normally acceptable to those of the left and right wings of politics alike. Political power must be bridled, but the constructive power of government must also be accepted, and strategy 2 and the path of stronger *political* regulation, rather than ethics-from-withining and self-regulation is here advocated as necessary, if not sufficient, to solve some of the foundational problems generated by AI.

While working from within the system may allow the ethicist to slightly change the direction of company strategies and policies, also working from without seem to allow for a larger canvas on which to sketch new and fundamentally different solutions. Solutions in which 'humane', 'responsible', and 'trustworthy' technology is not determined by the logic of the industries in power, but by drawing on more fundamental political and moral philosophy and fundamental theories of disruptive innovation and social change.

In this light, more work needs to be done on the relationships between power, social change, and technology, including on a more relational conception of power. Power is not only dangerous; it is also necessary to change the world. Moreover, one may not only take a pluralist approach towards ethical theories but also towards the strategies discussed here. Since clearly both strategies have advantages, a combination might be desirable to gain both of these advantages and mitigate some of the disadvantages discussed.

Declarations

References

1. Sætra, H.S.: Big Data's Threat to Liberty. Academic Press, London (2021)
2. Spohr, D.: Fake news and ideological polarization: Filter bubbles and selective exposure on social media. Bus. Inf. Rev. **34**(3), 150–160 (2017)
3. Pariser, E.: The Filter Bubble: What the Internet is Hiding from You. Penguin, München (2011)
4. Sunstein, C.R.: # Republic: Divided Democracy in the Age of Social Media. Princeton University Press, Princeton (2018)
5. Carr, N.: The Shallows: What the Internet is Doing to Our Brains. WW Norton & Company, New York (2020)
6. Noble, S.U.: Algorithms of Oppression: How Search Engines Reinforce Racism. New York University Press, New York (2018)
7. Van de Poel, I., Royakkers, L.: The ethical cycle. J. Bus. Ethics **71**(1), 1–13 (2007)
8. Müller, V.C.: Ethics of artificial intelligence and robotics. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy (Summer 2020 Edition). Springer, Berlin (2020)
9. Coeckelbergh, M.: AI Ethics. MIT Press, London (2020)
10. Culpepper, P.D., Thelen, K.: Are we all amazon primed? Consumers and the politics of platform power. Comp. Pol. Stud. **53**(2), 288–318 (2020). https://doi.org/10.1177/0010414019852687
11. Gillespie, T.: The politics of 'platforms.' New Media Soc. **12**(3), 347–364 (2010). https://doi.org/10.1177/1461444809342738
12. Sagers, C.: Antitrust and tech monopoly: A general introduction to competition problems in big data platforms: testimony before the committee on the judiciary of the ohio senate (October 17, 2019). https://doi.org/10.2139/ssrn.3471823
13. Turkle, S.: Alone Together: Why we Expect More from Technology and Less from Each Other. Hachette, Paris (2017)
14. Sætra, H.S.: The tyranny of perceived opinion: Freedom and information in the era of big data. Technol. Soc. **59**, 101155 (2019). https://doi.org/10.1016/j.techsoc.2019.101155
15. Véliz, C.: Privacy is Power. Bantam Press, London (2020)
16. Solove, D.J.: Privacy and power: Computer databases and metaphors for information privacy. Stanf. Law Rev. **53**, 1393 (2000)
17. Yeung, K.: 'Hypernudge': Big data as a mode of regulation by design. Inf. Commun. Soc. **20**(1), 118–136 (2017). https://doi.org/10.1080/1369118X.2016.1186713
18. Sætra, H.S.: When nudge comes to shove: Liberty and nudging in the era of big data. Technol. Soc. **59**, 101130 (2019). https://doi.org/10.1016/j.techsoc.2019.04.006
19. Buolamwini, J., Gebru, T.: Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency, pp. 77–91 (2018)
20 Ebell, C., et al.: Towards intellectual freedom in an AI Ethics Global Community. AI Ethics **1**, 131–138 (2021)
21. Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: Can language models be too big. Proc. FAccT (2021). https://doi.org/10.1145/3442188.3445922
22. Brevini, B.: Black boxes, not green: Mythologizing artificial intelligence and omitting the environment. Big Data Soc. **7**(2), 2053951720935141 (2020). https://doi.org/10.1177/2053951720935141
23 van Wynsberghe, A.: Sustainable AI: AI for sustainability and the sustainability of AI. AI Ethics (2021). https://doi.org/10.1007/s43681-021-00043-6
24. Mueller, D.C.: Public Choice III. Cambridge University Press, Cambridge (2003)
25. Downs, A.: An economic theory of democracy. Harper & Row, New York (1957)
26. Walker, K., Wan, F.: The harm of symbolic actions and greenwashing: Corporate actions and communications on environmental performance and their financial implications. J. Bus. Ethics **109**(2), 227–242 (2012)
27. Herrman, J.: "We're Stuck with the Tech Giants. But They're Stuck With Each Other. New York Times Magazine, New York (2019)
28. Sen, C.: The "Big Five" Could Destroy the Tech Ecosystem. Bloomberg, New York (2017)
29. Petit, N.: Big Tech and the Digital Economy: The Moligopoly Scenario, p. 11. Oxford University Press, Oxford (2020)
30. Foer, F.: World Without Mind. Random House, New York (2017)
31. Brown, S.D.: Michel Serres: Science, translation and the logic of the parasite. Theory Cult. Soc. **19**(3), 1–27 (2002)
32. Serres, M.: The Parasite. University of Minnesota Press, Minneapolis (2013)
33. de Jong, S.: Constructive complicity enacted? The reflections of women NGO and IGO workers on their practices. J. Intercult. Stud. **30**(4), 387–402 (2009)
34 Spivak, G.C.: A Critique of Postcolonial Reason: Toward a History of the Vanishing Present. Harvard University Press, Cambridge (1999)
35. Baring, E.: Liberalism and the Algerian war: The case of Jacques Derrida. Crit. Inq. **36**(2), 239–261 (2010)
36. Bergen, M., Brustein, J.: Google Protest Leader Leaves, Warns of Company's Unchecked Power. Bloomberg, New York (2019)
37. Hao, K.: "I Started Crying": Inside Timnit Gebru's Last Days at Google—and What Happens next. MIT Technology Review, London (2020)
38. Lorde, A.: The Master's Tools will Never Dismantle the Master's House. Penguin, London (2018)
39 Næss, A.: Ecology, Community and Lifestyle: Outline of an Ecosophy. Cambridge University Press, Cambridge (1989)
40 Vallor, S.: Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. Oxford University Press, Oxford (2016)
41. Nolt, J.: Environmental Ethics for the Long Term: An Introduction. Routledge, New York (2014)
42. Rachels, J., Rachels, S.: The Right Thing to Do: Basic Readings in Moral Philosophy. McGraw-Hill, New York (2015)
43. Sætra, H.S.: Privacy as an aggregate public good. Technol. Soc. **63**, 101422 (2020). https://doi.org/10.1016/j.techsoc.2020.101422
44. Corbett, E.P.: Classical rhetoric for the modern student. Oxford University Press, New York (1990)
45. Abizadeh, A.: The grammar of social power: power-to, power-with, power-despite and power-over. Polit. Stud. (2021). https://doi.org/10.1177/0032321721996941
46. Sattarov, F.: Power and Technology: A Philosophical and Ethical Analysis. Rowman & Littlefield, Lanham (2019)
47. Olson, M.: The Logic of Collective Action: Public Goods and the Theory of Groups, Second Printing with a New Preface and Appendix. Harvard University Press, Cambridge (1971)
48. Aiken, S.F.: The significance of Al Gore's purported hypocrisy. Environmental Ethics **31**(1), 111–112 (2009)
49. Lanier, J.: Ten Arguments for Deleting your Social Media Accounts Right Now. Random House, New York (2018)

50. Wenz, P.S.: Minimal, moderate, and extreme moral pluralism. Environ. Ethics **15**(1), 61–74 (1993)

51. Dewey, J.: Three independent factors in morals. In: Boydston, J.A. (ed.) The Later Works, 1925–1953, 5th edn., p. 1930. Southern Illinois University Press, Carbondale (1981)

52. Merton, R.K.: Science and technology in a democratic order. J. Legal Polit. Sociol. **1**(1), 115–126 (1942)

53. Aristotle, A.: Rhetoric. Hackett Publishing Company, Inc., Indianapolis (2018)

54. Bietti, E.: From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 210–219 (2020)

55. Næss, A.: Økologi, Samfunn, Livisstil. Bokkubbens Kulturbibliotek, Oslo (1999)

56. Sætra, H.S., Fosch-Villaronga, E.: Research in AI has Implications for Society: How do we Respond? Morals Mach. **1**(1), 60–73 (2021)

57. Zuboff, S.: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019. PublicAffairs, New York (2019)

58. Boyd, D., Crawford, K.: Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Inf. Commun. Soc. **15**(5), 662–679 (2012). https://doi.org/10.1080/1369118X.2012.678878

59. Prabhu, V.U., Birhane A.: Large image datasets: A pyrrhic win for computer vision? arXiv preprint arXiv:2006.16923 (2020)

60. Marcuse, H.: One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Routledge, New York (2013)

61. Chomanski, B.: The missing ingredient in the case for regulating Big Tech. Minds Mach. (2021). https://doi.org/10.1007/s11023-021-09562-x