Towards a Human-Centred View on Digital Technologies

Hannah Trittin-Ulbrich and Kirsten Martin

Introduction

The ethical concerns emerging from the proliferation of digital technologies have attracted attention from not only the popular press, but also a growing number of scholars across disciplines. This is not surprising, given that technology has always been met with ethical and critical examination. Arguably, the printing press and pony express were both viewed as disruptive innovations with value-laden design decisions and ethical implications. The steam engine was seen as an abomination against the gods by combining water and fire. And, the critical scholarly examination of technologies has included bicycles, plastics, seatbelts, bridges, among many other technologies (e.g. Winner, 1980). Even within our more current innovations, the ethical examination of digital technologies has advanced within engineering and philosophy for decades (e.g. Johnson, 1985).

However, we, that is, business ethicists, should not leave the interrogation of ethical concerns of digital technologies to ethicists of other disciplines (i.e. AI/artificial intelligence ethics scholars). In this essay, we argue that our discipline is uniquely equipped to interrogate the ethical implications of digital technologies for business and society. We use the term digital technologies to mean information and communications technologies that rely on the latest data analytics techniques to include a range of technologies, e.g. artificial intelligence (AI), social media, platforms, facial recognition or blockchain. Digital ethics, defined here as current efforts to discern the ethics of, and corporate responsibilities for, digital technologies, provide important counter-perspectives to the tech-hype that is fueled by the "Internet-industrial" complex (Flyverbom et al., 2019), that is, those private, public and other actors involved in the development and governance of the Internet and digital technologies.

With this essay, we hope to inspire future business ethics research to further interrogate what constitutes a human-centred approach to the development and application of digital technologies in the business context. We propose that a critical, human-centred approach implies that digital technologies should be developed and adopted in the interest, and to the benefit, of those individuals who are affected by them. To that end, importantly, we must avoid falling into the trap of (unintentionally) subscribing to imperative arguments regarding the inscrutability, efficiency and profitability of digital technologies. Let us explain.

What Business Ethicists Should Avoid When Dealing with Digital Technologies

First, digital technologies are often sold with the false claim of *inscrutability*. Artificial intelligence, in particular, has often been debated as working in a "mysterious", autonomous way. In view of such autonomous decision making of artificial intelligence, firms and their representatives claim that they can no longer be held accountable for the impact that such technology produces. Automated decision making is also suggested as creating fairer, more objective outcomes than human beings. And yet, if an algorithm is found to create wrongful or even harmful outcomes, firms tend to displace responsibility towards the autonomous artificial agent: "It's not us, it's them". However, digital technologies designed to be inscrutable are more about corporate power than any design requirement (Kroll, 2018; Pasquale, 2015). Engineers have developed ways to test and report ethical issues of AI and even machine learning—that is, the design decision to make a programme inscrutable is a decision and should not be taken as a given (Martin, 2019).

This brings us to the second fallacy of digital technologies: digital technologies also (falsely) promise *efficiency* and the hyper-rationalization of firm activities. Goodbye slow, flawed human decision making, welcome rational and efficient automated decision making! Brave new digital world. Arguably, such narration relies on two flawed assumptions: (1) a limited, economic theory of the firm that conceives of the business firm as a purely economic actor, whose only goal must be the enlargement of the shareholder value through continuous enhancement of the firm's efficiency; and (2) a view of technology development as neutral and objective, and devoid of value-laden decisions made throughout the design and development process (Martin, 2022).

Finally, digital technologies are also falsely hailed as suitable means to increase corporate *profitability*. We are told that digital technologies allow businesses to "move fast and break things" (as famously proclaimed as the key to success of the social networking platform Facebook by CEO [Chief Executive Officer] Mark Zuckerberg), that is, to disrupt existing markets and industries, allowing them to quickly enlarge their value and attractiveness for venture capital. In this sense, the efficiency of digital technologies is promised to directly result in a firm's increase in profitability.

Questioning these three common assumptions of inscrutability, efficiency and profitability is not enough in pursuing a more critical, human-centred approach to digital technologies. We therefore now turn to identifying three unique avenues of scholarship on the ethics of digital technologies where business ethics research has a unique grounding and

perspective that allow for a human-centred view on digital technologies.

Where Business Ethics Should Engage

Who Should be Held Accountable for the Impact of Digital Technologies?

First, in order to move business ethics research forward, we must avoid falling into the trap of thinking that businesses cannot be held accountable for the moral implications of the digital technologies they use or produce. Arguably, decision making augmented with AI often takes place behind closed (corporate) doors, hiding from public scrutiny and oversight. However, business ethics research is essentially the study of accountability: that is, the field questions who is responsible for an action or outcome under which premises, as well as providing reason as to *why* firms are responsible for their decisions and their impact on society.

Mistakenly assuming that digital technologies provide more efficient, accurate decisions, and are outside the realm of any critical examination or moral evaluation leads scholars to incorrectly see the development of digital technologies as being deterministic and outside their scope. Judging technologies on efficiency and treating digital technologies as inscrutable products also shields corporations from being held accountable for the value-laden decisions made in the design, development and deployment of algorithms. At the other end of the spectrum, pretending that digital technologies are only as ethical as how society uses them—as if the design decisions have no bearing on the moral implications of their use—allows firms who design and develop digital technologies to avoid the sharp gaze of critical theorists who wish to hold them responsible for their decisions.

To lift the "veil of the technological imperative" (Martin, 2022) and critically examine the moral implications of the design, development and use decisions around digital technologies, business ethicists should raise questions in regard to who can be held accountable for how digital technologies impact business and society, as well as to why this is so. In particular, we must further interrogate the design and development process of digital technologies, and must ask how we can hold those actors, including information technology experts or software developers, accountable for their decisions in this process. Business ethics scholars should follow others who acknowledge that digital technologies have biases that are value-laden (Friedman & Nissenbaum, 1996; Johnson, 2004) or have political dimensions (Winner, 1980), while also identifying how individuals, corporations and society can control that same technology.

Digital technologies have value-laden biases along several dimensions which can serve as avenues for future normative work within business ethics (Martin, 2022).

In particular, digital technologies are biased towards and designed for a preferred set of actions that create (or destroy) value for stakeholders, uphold (or violate) ethical principles and reinforce (or undermine) stakeholder rights and dignity. In this way, digital technologies can be seen as embodying corporate policy or the norms and rules of the organization which are then enacted into decisions. As such, business ethics should expand the ethical evaluation of digital technologies to include the value-laden decisions made around the outcome of digital technologies, the criteria for whether a technology *works*, the choice of data used as well as the assumptions made in the development of these technologies.

How do we Theorize the Goal of the Firm in Relation to Digital Technologies?

Business ethics research is also uniquely equipped to counter the imperative arguments regarding the efficiency and profitability of digital technologies. Our field has a longstanding tradition of questioning one-sided and purely functionalist constructions of the goal of the business firm and, therefore, should naturally respond with caution to idealizations of the automated firm.

In business ethics, we argue (regularly) that the focus on shareholder value is mistakenly socially constructed as the only goal a firm should pursue. Similarly, the idea of efficiency as the only goal of digital technologies is socially constructed and inherently value based. Indeed, efficiency is usually constructed to serve only specific sets of actors—the firm and its shareholders—without consideration of other actors that have a stake in the firm's activities and involvement with digital technologies. For example, for whom is the hiring AI programme efficient? How are the goals of the firm served if an AI programme to read resumes or assess interviews consistently makes mistakes, but does so "efficiently"? Business ethicists aiming to advance our field in the study of the ethics of digital technologies should carefully interrogate and expose the connection between the development and use of digital technologies that can and should be aligned with a thicker conception of the goal of the firm. Much of the current work on the ethics of digital technologies relies on a thin, shareholder wealth maximizing view of the firm.

The case of social media content moderation algorithms—designed to quickly promote *lawful but awful* material in order to increase user engagement—exemplifies how digital technologies may serve a specific goal while being destructive to firm value from the perspective of society at large. Within the field of AI ethics, Thomas & Uminsky (2020) call for multifaceted outcomes for measuring the effectiveness of data analytics programmes and even the danger of allowing an analytics programme's outcome variable, which is being optimized, to dominate the decision

making of the firm. Social media's fixation on user engagement, to the detriment of all other measures, exemplifies this danger.

Broadening notions of the goal of the firm also requires broadening conceptions of the value of digital technologies. Many proponents of the digital economy try hard to promote a dichotomous concept of digital technologies as only productive if unconstrained by governmental regulation around fairness. To contribute to a more humane approach to digital technologies, business ethics scholars should not approach fairness and efficiency as opposing, but rather as complementary, goals in the application of digital technologies. We must critically interrogate when, and under which conditions, businesses can achieve both a fair and productive application of digital technologies, and thus, also contribute to fairer and productive businesses. We must ask, how can a corporation live up to its corporate digital responsibility? Moreover, how do societal demands shift and change in regard to corporate digital responsibilities, and what implications emerge from changing and evolving societal expectations regarding fair and responsible business conduct for the application of digital technologies? Who are the stakeholders of digital technologies, and how should technologies be designed and developed with these stakeholders in mind?

Finally, to contribute to a humanistic approach to digital technologies, business ethicists should further interrogate the question who the stakeholders of digital technologies are and how technologies should be designed and developed with these stakeholders in mind. In the past, business ethics research has provided much insight into how and why businesses should respect the interests and rights of various stakeholders. Following this line of reasoning, business ethics scholars should interrogate how the business application of digital technologies enhances both the firm's performance and contributes to serving the interests of all stakeholders.

In business ethics, stakeholders are generally considered those individuals that are affected by or can be influenced by a business decision or action. The general assumption here is that stakeholders engage knowingly and often voluntarily with businesses and have some sort of formal relationship with the firm, for example, by being their customers, employees, suppliers (Freeman, 1984). Normally, to categorize stakeholders and prioritize their interests, we distinguish primary and secondary stakeholders, and group them into a "manageable" stakeholder community that a business regularly engages with. Stakeholders are voluntary, we assume, and in a relationship with the firm.

However, with digital technologies, our current approach as to who constitutes a stakeholder is challenged. Not only are decisions augmented with AI hidden from market governance or public oversight, but digital technologies impact actors that have no formal relationship with the firm and are not voluntary. Social media content moderation algorithms impact not only the advertisers, which are customers and serve as revenue sources, but also users and even individuals and groups not on social media. The recommendation of hate groups by algorithms or violence against dissidents impacts not only the users of social media, but the targets of these violent groups who are not on social media. The original definition of a firm's stakeholder—those who are influenced by or influence the firm (Freeman, 1984)—is a closer approximation to the issues faced today with digital technologies. Such a definition does not require stakeholders to be voluntary, nor in an immediate relationship with the firm. It is therefore crucial that ethicists begin by reviewing who the stakeholders of digital technologies are and whose voices should be considered when developing them.

Based on the original definition of stakeholder (Freeman, 1984), we see three types of unaware and often overlooked stakeholders, where business ethics research could leverage our theories to understand how firms should manage these stakeholder relationships. First, stakeholder groups of digital technologies may include, for example, "unaware" stakeholders who are being impacted by a digital technology but are not aware of the digital technology being used. Unaware stakeholders include individuals whose job application, social networking or dating site pictures on the Internet are used to train face recognition technology. Similarly, this type of stakeholders includes those confronted with the results of automated decision making without realizing that such technology is applied. Since some technologies may crawl and detect data points in the global Internet, this makes any Internet user potentially a part of this stakeholder group—reducing the very idea of a more or less stable and controllable "stakeholder community" to absurdum.

Second, while arguably, many relations businesses traditionally hold with their stakeholders are slanted in terms of distribution of power or information, digital technologies create new "unequal" stakeholders that interact with businesses in the form of unequal relations. These include "gig" or platform economy workers that often work at the whim of an algorithm, gaining work assignments and being evaluated through opaque forms of algorithmic management and control. Similarly, platform users, including social media users, have limited insights into what algorithmic decision has led to certain content being shown to them, while other remains hidden. Even businesses may find it hard to gain insights into why e-commerce platforms decide to show their products only to certain customers, reducing even large business conglomerates to an "unequal" business partner to these platform titans.

Third and finally, digital technologies also create new "invisible" stakeholders, that is stakeholders that are invisible to the digital economy business models. Invisible from the public eye, and often hidden in plain sight from scholarly inquiry, masses of poorly paid independent contractors from

the global south "curate" the content that is published on social media platforms by reviewing and deleting masses of disturbing and often downright illegal data. Other stakeholder groups of this category also include those factory workers working in the delivery centrs of large e-commerce retailers.

These stakeholders are, within our current parlance, legitimate but marginalized. Future business ethics research concerned with digital technologies should interrogate how businesses should have an obligation for the impact their technologies have on these and other stakeholders. How should digital technologies be designed to give unaware and/or silent stakeholders a voice? Which role can artificial intelligence play in creating new forms of automated accountability? How can the rights of those being unaware of their status as being a stakeholder be upheld? How can more transparent and fair working conditions be ensured in platforms, and how can workers' dignity and rights be secured in fully automated work arrangements that lack governmental regulation and public oversight? Providing answers to these questions will contribute to a more critical, human-centred approach to digital technologies.

Conclusion

With this essay, we hope to inspire future business ethics research to further interrogate what constitutes a human-centred development and application of digital technologies in the business context. The proliferation of digital technologies promises great human advancement, while also raising questions regarding the ethical and responsible development and application of these technologies by businesses. We welcome the growing number of business ethicists paying attention to the critical, problematic and "dark" implications of these technologies for business and society (e.g. Trittin-Ulbrich et al., 2021). We have outlined imperative arguments regarding the inscrutability, efficiency and profitability of these technologies and we have outlined three areas of growth for future research through which business ethics scholars can contribute to a more critical and human-centred approach to digital technologies. We are looking forward to seeing their efforts!

Reimagine Corporate Social Responsibility in the Age of Artificial Intelligence

Shuili Du

Artificial Intelligence and its Double-Edged Effects

As artificial intelligence (AI) increasingly permeates the business world and modern society, companies need to rethink and broaden the scope of their corporate social responsibility strategies and initiatives to deal with key ethical and socio-technical issues triggered by AI and related technologies. Defined as "the ability of machines to carry out tasks by displaying intelligent, human-like behavior" (e.g. machine learning, computer vision, speech recognition and natural language processing: Russell & Norvig, 2016), AI is transforming our economy. The global AI market size is forecast to grow from $58.3 billion in 2021 to $309.6 billion by 2026, at a compound annual growth rate of 39.7% (Markets & Markets, 2021). AI technologies are being deployed in diverse sectors, ranging from finance, health care and transportation, to national security, criminal justice and smart cities, augmenting human capabilities in significant ways and making a profound impact on the world. However, AI is a mixed blessing. On the one hand, it promises scientific breakthroughs and advancement of humanity with its superior processing speed, limitless recall and self-improving learning ability. On the other hand, it is fraught with a host of unprecedented ethical and socio-technical challenges, such as AI algorithmic biases, machine ethics, data privacy, job replacement by AI and exacerbated digital inequity.

AI follows the trajectory of exponential growth, and it seems that our society is marching inexorably towards artificial superintelligence—the point of singularity—when AI systems will be self-aware and outperform humans in nearly all areas (Bostrom, 2014). Super-intelligent AI will be capable of complex goal setting and can engage in scientific discovery and artistic creativity (Tegmark, 2017). Such systems hold enormous promise in transforming every aspect of our society for the better by, for example, repairing damage done to the natural world and eradicating poverty and diseases. At the same time, when machine intelligence eclipses human intelligence, technological growth becomes uncontrollable and irreversible, resulting in unforeseeable changes to human civilization. AI is humanity's biggest existential threat, as Elon Musk famously stated.

The future as increasingly mediated by AI is both fascinating and terrifying. Corporate social responsibility (CSR) scholars can play a big role in shaping the short-term and long-term future of ethical and socially responsible AI. To embrace the power of AI while minimizing its downsides, companies should reimagine their CSR strategies and practices to turn the unique social challenges of AI into business opportunities. In the short term, businesses need to tackle an array of ethical and socio-technical issues surrounding AI nowadays, including AI biases, machine ethics, data privacy, cybersecurity, individual autonomy, job replacement by