



The Business School  
for the World®

**Working Paper**

2021/23/ATL

## **Corporate Responsibility Meets the Digital Economy**

Leena Lankoski

Aalto University School of Business and INSEAD, [leena.lankoski@aalto.fi](mailto:leena.lankoski@aalto.fi)

N. Craig Smith

INSEAD, [craig.smith@insead.edu](mailto:craig.smith@insead.edu)

We examine how digital transformation of the economy may affect the field of corporate responsibility (CR) and whether some of its foundational assumptions need to be re-examined. Grounding our analysis in the relevant theory, we identify five key digital economic phenomena from a CR-relevant perspective, outline their associated ethical opportunities and threats, and map those opportunities and threats against three core questions that define the domain of the CR field: responsible for what, responsible towards whom, and who is responsible? We find that while there are foreseeable changes required of the field in relation to all three questions, it is the third one that poses the biggest challenges, questioning foundational assumptions on the allocation of responsibility (notably, that responsible entities are clearly identifiable and are human agents only). We develop a research agenda to address the concerns raised and conclude with a discussion of our contribution and its implications.

Keywords: Artificial Intelligence; Big Data; Business Ethics; Corporate Responsibility; Digital Economy

Citation style: Academy of Management Journal

Electronic copy available at: <http://ssrn.com/abstract=3845940>

Working Paper is the author's intellectual property. It is intended as a means to promote research to interested readers. Its content should not be copied or hosted on any server without written permission from [publications.fb@insead.edu](mailto:publications.fb@insead.edu)

Find more INSEAD papers at <https://www.insead.edu/faculty-research/research>

Copyright © 2021 INSEAD

Digital technology is transforming the economy, bringing with it tremendous gains for society but also significant challenges. The phrase “everything that can be digital will be” is often quoted, and the transformations brought by digitalization are described with words such as ‘sea change’ (Calo, 2014), ‘revolutionary’ and ‘disruptive’ (Andriole, 2017). The changes have been predicted to be so deep that even the foundations of entire fields need to be revisited. Indeed, such research is beginning to emerge: there are now studies about how digitalization affects the fields of economics (Parkes & Wellman, 2015), strategy (Constantiou & Kallinikos, 2015), management information systems (Lyytinen & Grover, 2017), accounting (Arnaboldi, Busco & Cuganesan, 2017), and working and organizing (Faraj, Pachidi, & Sayegh, 2018). These studies highlight how digitalization calls for a re-examination of the continued applicability of previous assumptions to bring a field to par with contemporary developments.

Such an examination is notably absent in relation to the field of corporate responsibility (CR). There is a large and growing body of literature examining individual ethical issues in the digital economy – an increasingly hot topic not only in academic research but also in policy-making and in corporate practice as the ethical implications of digitalization are becoming more evident. There are also more holistic discussions of corporate responsibility and digitalization (e.g., Flyverbom, Deibert, & Matten, 2019). Still, these studies have not addressed the question of how CR as a field, including its basic foundations and assumptions, may be affected by digitalization. Yet, a logical consequence of a digital disruption in the economy would be a corresponding disruption in CR. This, however, depends on the nature of the current understanding of CR and its robustness in the face of the digitalization challenge; in other words, the extent to which it can accommodate changes brought by digitalization without having to change any fundamental

assumptions. Thus, the question begging examination is: How does digitalization change the field of CR, how fundamentally, and what remains unchanged?

To address this gap, we provide the first systematic examination of digital economy and CR from this novel ‘big picture’ perspective. We ask: How may the fundamental changes in the digitalizing economy shape the field of CR and require scholars and managers to think differently about CR? To answer this question, we proceed in three steps. First, from a stakeholder-theoretic perspective, we dissect digitalization in the economy into five CR-relevant key phenomena and discuss their associated ethical opportunities and threats. Second, we delineate the domain of the CR field (specifically, as regards the assignment and adoption of responsibilities) through three core questions and examine how those ethical opportunities and threats may affect the answers to the questions. Third, we choose the area with the largest potential for fundamental changes for a deeper examination and develop a research agenda to address that area. We conclude the paper with a discussion of our contribution and its implications for research and practice.

## FROM DIGITAL TECHNOLOGY THROUGH DIGITAL ECONOMY TO ETHICAL ISSUES

New digital economy phenomena are unleashed by developments in digital technology, especially the coming together of big data and artificial intelligence. The term ‘big data’ generally refers to the accumulation of large data masses in a digital environment. Aside from huge volumes of data, what differs is also their variety, velocity and granularity (Martin, 2015). Artificial Intelligence (AI) generally refers to algorithms that undertake problem-solving tasks. Thanks to machine learning, rather than being directly programmed by humans, machines can observe and learn from data given to them for training purposes or from past experience and generate the algorithms

needed to achieve a goal. The terms often overlap, as references to big data can include the analysis of such data (e.g., Martin, 2016), while AI feeds on big data.

### *Identifying Digital Economy Phenomena from a CR-relevant Perspective*

There are different ways of identifying digital economy phenomena enabled by digital technology. For our purposes, we need to examine them from a CR-relevant perspective. A central theory underlying CR is stakeholder theory (e.g., Freeman, Harrison, Wicks, Parmar, & de Colle, 2010). It identifies the stakeholder relationship as the most useful unit of analysis (Freeman, Phillips, & Sisodia, 2020). Thus, we base our identification of the key digital economy phenomena on how stakeholder relationships are affected: how firms operate in novel ways towards their stakeholders, and how stakeholders operate in novel ways among themselves and towards the firm. Within these categories we distinguish five (mutually non-exclusive) CR-relevant phenomena in the digital economy (see Figure 1). Although our capturing of the digital economy through these phenomena is new, each of the individual phenomena are already well-established in previous literature.

---

Insert Figure 1 about here

---

First, it is possible for a firm to employ digital technology in its various functions and towards its various stakeholders via (i) *digital marketing*, (ii) *algorithmic management*, and (iii) *autonomous processes in its products and services*. Second, digital technology also enables the stakeholders of the firm (jointly with the firm) (iv) to interact among themselves as participants to the *sharing economy*, enabled by the platforms set up by firms, and (v) to interact with the firm as a result of *increased transparency and enhanced stakeholder governance*. Next we describe

shortly these digital economy phenomena and the kinds of ethical opportunities and threats that have been identified in connection with them (see Table 1).

### *Key Digital Economy Phenomena and their Ethical Implications*

#### Digital Marketing

The term digital marketing captures a wide variety of measures used by firms in an attempt to market their products and services to customers more effectively with the help of digital media. Digital marketing may occur through the Internet (web pages, social media platforms), mobile phones (SMS, MMS, apps), e-mail, and intelligent displays; in e-commerce the entire transaction itself is digital. Digital marketing can involve the personalised targeting of the timing, content, and form of messages, as well as the dynamic and personalized setting of prices (Seele, Dierksmeier, Hofstetter, & Schultz, in press). All this may be based on the customer's previous purchasing behaviour, online behaviour (e.g., on the firm's web page or online shop or on other websites, on social media platforms such as Twitter and Facebook, or in search engines), the customer's location, or even on his or her current moods obtained through facial coding. Digital marketing thus relies heavily on data about the customer and on algorithms to make sense of these data (West, 2019).

Calo (2014) summarizes many of the ethical issues relating to digital marketing in his discussion of 'digital market manipulation'. Market manipulation refers to firms using "what they know about human psychology to set prices, draft contracts, minimize perceptions of danger or risk, and otherwise attempt to extract as much rent as possible from their consumers" (Ibid., p. 1001). In a digital context, the availability of big data creates an information asymmetry which dramatically enables firms to influence the consumers, even at an individual level. As a result,

consumers are less able to pursue their self-interest through their choices and may be led to purchasing products and services they did not need or want or to paying higher prices. Thus, digital market manipulation can result in economic harm, privacy harm (subjective and objective), and autonomy harm for the consumer (Ibid.), though it should also be noted that there may be benefits too, such as better understanding of individual consumer price elasticity of demand, bringing benefits to some consumers in the form of lower prices when companies price discriminate. In some ways, these are old concerns about marketing in a digital guise. However, the scope to exploit consumer psychology at massive scale and in depth is unprecedented.

### Algorithmic Management

In addition to digitalization transforming relationships between firms and their customers, it also affects the internal management of organizations (Schildt, 2017). Algorithmic, data-driven management (Lee, Kusbit, Metsky, & Dabbish, 2015) is emerging both within firms, where algorithms track the performance of employees to optimize organizational structures and human resource management (see Leicht-Deobald, et al., 2019), and between firms, where digital supply chain management may be applied to contractors (see Schildt, 2017). It can be used to understand not only the behaviour and performance of individual employees but also broader employee network dynamics and to coordinate work tasks (Baesens, De Winne, & Sels, 2017). When algorithmic management is taken to the extreme, “the core function of management has gone from managing the business to managing the bots that are managing the business” (O’Reilly, 2016).

When it comes to employee monitoring and control, the ethical issues revolve around questions that include freedom, privacy, and respect (Martin & Freeman, 2003). These issues arise because algorithmic management makes it possible to capture and evaluate employee performance

in real time and at a detailed level to the extent that questions of surveillance clearly arise – it has been called “Taylorism on steroids” (O’Connor, 2016). For instance, the algorithm of the takeaway delivery company Deliveroo systematically monitors all aspects of the tasks of its couriers (time to accept orders, travel time to restaurant and to customer, time at customer, late orders, unassigned orders) and compares these against a benchmark (Ibid.). On the positive side, algorithmic management of employees may guarantee a certain consistency: an algorithm has no good or bad days, nor does it resort to favouritism (although it might be biased in ways that favour particular employees).

#### Autonomous Processes in Products and Services

A third type of digital economy phenomena is the emergence of artificial agents that operate autonomously. Such artificial agents can be software programs (bots) making decisions, or hardware-software combinations (robots) carrying out physical activities, and they “perform tasks on behalf of humans and do so without immediate, direct human control or intervention” (Johnson, 2015, p. 708). According to Hellström (2013), an often referenced definition goes that an autonomous agent is “a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future.” Firms can employ artificial agents in their services (e.g., in credit decisions), or the firm’s product can be an artificial agent (like a robot vacuum cleaner or a self-operating vehicle).

Autonomous agents are made possible through digital technology and AI in particular. Consequently, the relevant ethical issues are indeed those that have been presented as ethical concerns with AI. One such concern relates to the fact that machines are different from humans in their reasoning and operation. Machines lack human moral considerations even though their

decisions can have moral consequences. Whereas human action is generally restricted by a moral conscience (albeit to a different degree for different individuals), machines could undertake immoral or illegal actions without the blink of a (mechanical) eye. This is, unless they are programmed or taught to incorporate human ethical reasoning. Ironically, however, another concern relates to the fact that machines might be similar to humans: if they learn by imitating human behaviour, they will perpetuate the same biases and repeat the same unethical habits and behaviours as humans. In any case, algorithmic accountability or explainability is a relevant issue because as AI gets complex and self-improving, its internal workings become a 'black box' and it is not possible even for experts to know why the algorithm has reached a particular result. A final concern relates to AI replacing humans. In the more immediate future, there is the question of job loss as artificial agents are occupying jobs until now conducted by human workers; but then there is the ultimate question of technological singularity which refers to artificial superintelligence surpassing human intelligence and potentially taking control of itself and our societies.

On the other hand, assuming that it were possible to incorporate ethical reasoning into AI so that it would be accompanied by artificial morality, machines might make *more* ethical decisions than humans. There are two reasons for this. One is that with humans, the distribution of morality is wide but machines are all just as ethical as they are designed to be. Another is that humans have weaknesses like cognitive limits and biases, visceral reactions and a lack of self-control and are therefore not always able to choose the ethical course of action; machines lack such limitations.

Sharing Economy



Still another manifestation of the digital economy is the emergence of the so-called sharing economy. It is made possible by digital platforms provided by firms where algorithms connect product/service providers with product/service recipients, and which also include a rating system to build trust between the parties. Four features of sharing economy organizations can be identified (Gerwe & Silva, 2020): they are organized as digital platforms, they facilitate peer-to-peer transactions, they emphasize temporary access rather than ownership, and they are focused on the use of underutilized capacity. Thus, Gerwe and Silva define the sharing economy as “a socioeconomic system that allows peers to grant temporary access to their underutilized physical and human assets through online platforms.” Well-known examples of sharing economy companies include Uber and Lyft (in transportation), Airbnb (accommodation), and Amazon Mechanical Turk and TaskRabbit (tasks). Alternative terms have been proposed to describe different aspects of this phenomenon (e.g., ‘on-demand economy’, ‘peer-to-peer economy’) and it is useful to differentiate between businesses that are typically large and well-established professional providers and those with individuals as providers (which are more clearly peer-to-peer)—Uber and Airbnb have both.

There is the potential opportunity through the sharing economy to utilise excess capacity and save natural resources. The growth of vehicle sharing in various forms (e.g., Zipcar, Uber, BlaBlaCar) should in principle reduce demand for vehicle ownership. However, there are also ethical concerns, including questions of accountability and assignment of responsibility (Etter, Fieseler, & Whelan, 2019). Whether the platform-providing companies should be held accountable for the actions of the individuals (e.g., drivers) offering their services through those platforms has been debated from a legal as well as a CR perspective. Particularly noteworthy are responsibilities long-assumed by traditional providers in a sector that are avoided by new entrants: for example,

Airbnb has been criticized as unfair competition for hotels who must comply with lodging regulations and charge occupancy taxes, while car-share platforms like Uber have been accused of exploiting drivers who are considered independent contractors rather than employees (in most legal jurisdictions). Still another concern is that consumer-sourced rating systems might be discriminatory against certain protected groups; while companies are prevented from workplace discrimination by law, discriminatory bias could be present in the ratings given by individuals (Rosenblat, Levy, Barocas, & Hwang, 2017). Even the environmental benefits are questioned; some people may give up car ownership with the availability of ride-sharing options, for others these options may actually lead to a car purchase as it becomes a source of additional income, and in some cities lower prices for car-based transportation have reduced demand for public transportation.

#### Transparency and Stakeholder Governance

Finally, digital technology has enabled the rise in the economy of a multi-faceted phenomenon best described as transparency and stakeholder governance. There are more data available to the firm, including data about complex supply chains –which are also more granular and more real-time. This enables greater transparency and traceability should the firm wish to communicate this information to stakeholders. Furthermore, new digital media allow the firm to go beyond unidirectional communication to establish a two-way dialogue with stakeholders (Illia, Romenti, Rodriguez-Cánovas, Murtarelli, & Carroll, 2017), thus creating ‘dynamic transparency’ (Vaccaro & Madsen, 2009). On the other hand, digital media can also empower stakeholders such as activist networks or the general public to organize and communicate more effectively (see e.g., Illia et al., 2017) and hence influence and pressure organizations about their CR (Coombs & Holladay, 2015).

All this contributes to increasing transparency (even ‘super-transparency’, Austin & Upton, 2016, or making firms ‘naked’, Tapscott & Ticoll, 2003), stronger engagement of stakeholders in dialogue, and stronger governance of firms by stakeholders (Castelló, Etter, & Årup Nielsen, 2016; Etter, Colleoni, Illia, Meggiorin, & D’Eugenio, 2018).

Of the digital economy phenomena we have identified, transparency and stakeholder governance is the one that relies more on traditional information and communications technology than the latest developments in big data and AI (although the emerging blockchain technology will have major implications on the way that records are being kept and thus on traceability). For this reason perhaps, it is also the one that has been around the longest, and the scholarly literature that deals directly with CR and digitalization often tends to be from this perspective. This is also the phenomenon where the associated ethical issues are predominantly seen as opportunities. Transparency is considered important for stakeholders to be able to preserve their autonomy and make informed choices (Vaccaro & Madsen, 2009), and it is indeed one of the key principles of social responsibility (e.g., as recognized in ISO 26000). The ability of stakeholders to pressure firms to step up their CR efforts is also considered beneficial. On the other hand, swift dissemination methods could also be misused to spread false information (Vaccaro & Madsen, 2009), and the quality of information (e.g., in social media) may be more suspect (Wood, Mitchell, Agle, & Bryan, in press). As a consequence, stakeholders might exert pressure to steer companies in socially undesirable directions.

Table 1 summarizes the five digital economy phenomena with examples of the associated ethical issues.

---

Insert Table 1 about here

---

## DIGITAL ECONOMY AND THE CHANGING LANDSCAPE OF CR

In the preceding section we saw how developments in the digital economy lead to ethical issues, presenting both opportunities and threats. In this section, we analyse those ethical threats and opportunities to examine how they might affect the overall landscape of the field of CR. As a starting point, we describe the domain and boundaries of CR through three questions. Here we refer specifically to the assignment and adoption of responsibility. Whereas the field of CR can also be seen to contain strategic and operational management topics, it has an important normative core (Donaldson & Preston, 1995). We focus on this normative core and understand the landscape of CR to cover the questions of assignment of responsibility and the boundaries of this assignment. In our analysis we thus do not discuss questions that relate to the measurement and implementation of CR, or other such topics, even though they may also be affected by digitalization. Our intent is more fundamental.

### *Questions Defining the Landscape of CR*

In our initial account of the theoretical foundations of this paper, we observed that stakeholder theory is at the core of the field of corporate responsibility. As Donaldson and Preston (1995) explain in their seminal paper, while stakeholder theory is descriptive and instrumental, it is more essentially normative. By this they mean that “the normative basis for stakeholder theory involves its connection with more fundamental and better-accepted philosophical concepts” (1995, p. 81). Thus, consistent with theorizing in moral philosophy (e.g., Singer, 1994) and moral psychology (e.g., Gray & Wegner, 2009), we argue that the field of CR in relation to stakeholder theory may be analyzed through three key elements: moral agents (performing moral actions), moral patients (upon whom the actions are performed), and the moral actions themselves. These are distinct,

underlying dimensions to discussions of CR, and together they capture the scope of the assignment of corporate responsibility. First, moral actions are those which “can cause moral good or evil” (Floridi & Sanders, 2004, p. 364) and which may therefore create responsibilities for companies. The contemplation of moral actions in the context of CR thus results in the question of what can be considered moral actions or issues: *Responsible for what?* Second, moral patients consist of all entities that “can in principle qualify as receivers of moral action” (Ibid., p. 350). In the context of CR, this poses the question of who are the parties that may experience good or evil, or right or wrong, as a result of moral actions by companies, and towards which companies may therefore have responsibilities. This culminates in: *Responsible towards whom?* Third, moral agents are those entities that “can in principle qualify as sources of moral action” (Ibid., p. 349). From a CR perspective, this relates to the point of who are sources of moral action and who might therefore be assigned responsibilities for such actions: *Who is responsible?*

Thus, the field of CR may be described through three questions that delineate its domain and boundaries: responsible for what, responsible towards whom, and who is responsible? This set of questions is a collectively exhaustive description of the landscape, and the questions are also mutually exclusive. Below we elaborate on the interface between digitalization and CR through the lens of these three questions. For each question, we first describe established frames for thinking and then discuss how digitalization might change this thinking. We end by characterizing the nature of the potential changes from the perspective of the extent to which they challenge foundational assumptions in the CR field.

*Moral Actions: Responsible for What?*

The first of our analytical questions asks about the issues for which the firm might be responsible. There have been several attempts to capture the broad, diverse, and somewhat vague domain of CR by dividing it into a number of narrower issues. Among the most comprehensive and authoritative efforts to this end is the ISO 26000 standard, which identified social responsibility issues through an international multi-stakeholder process (ISO, 2010). It lists seven core social responsibility subjects, which are further divided into 36 more specific issues (not all of which will be relevant to every organization). Parallel and largely compatible issue lists can be found, for example, in the Global Reporting Initiative (GRI) reporting recommendations and their sector-specific supplements. The UN Sustainable Development Goals, adopted in 2015, also serve to pinpoint relevant domains for CR. Against all this, we can examine the changes brought by the digital economy.

Digital economy can, first of all, *make existing CR issues manifest themselves in novel ways*. For example, although AI has prompted questions about job replacement by robots, the fundamental social responsibility issue of employment is an old one. Same goes for working conditions, which arise anew in the ‘gig economy’. Or take the question of big data, digital marketing and privacy: even though it now comes up in the digital context, the underlying issue of consumer data protection and privacy was there before large-scale digitalization. In a similar vein, the fundamental social responsibility principle of transparency now manifests itself in a novel form with the ‘black box’ nature of some self-learning algorithms. However, while these are old issues in certain respects, they have gained new urgency.

Second, digitalization can *help alleviate or solve existing CR issues*. At a more technical level, improved transparency both within the firm and externally, brought about by the availability of more data and faster communications around such data, allows firms to track performance in

product chains and consumers to take informed action. This helps firms better implement their CR and also provides more incentives for doing so since reputation and legitimacy may be more readily at stake. More broadly, great expectations are being placed in the ability of AI in particular to help solve humanity's grand challenges, in areas such as energy consumption or medicine.

Third, the digital economy can *intensify existing CR issues*. The availability of big data makes many data-dependent activities in the consumer and employee domains possible in an unprecedented manner. Marketers have always collected information about customers (or firms about employees), but now this is happening on a different level. There is vastly more information; for instance, due to improved speech to text transcription, “all conversations that take place in organizations will soon be accessible to artificial intelligence” (Schildt, 2017, p. 26). It is more granular and originates from surprising sources, such as your robot vacuum cleaner selling your house map to marketers (Jones, 2017). It can be combined in imaginative ways and sophisticated inferences can be drawn from it by algorithms; Airbnb is said to use a “trait analyser” to detect dark personality traits such as narcissism by evaluating hundreds of signals with predictive analysis and machine learning (Simpson, 2020). As a consequence, marketers (or employers) will know more about you, including things you would not want to reveal (that you really need that flight for next April, no matter the price) – or things that you did not even yourself know about yourself, understanding of your “most personal motivations and vulnerabilities” (Jersinovic, 2017).

Fourth, the digital economy can *open up new CR issues*. One way is through intensification. Calo (2014) argues that digital market manipulation is not only a matter of degree but really different from ‘traditional’ market manipulation because, for the first time, personalization and intense systematization can be combined to enable targeting of individual consumers. This could be with welcome tailored content, though it could bring unwelcome content

and even leverage data against the consumer (e.g., in personalized pricing). At the time of writing this article, new CR issues could be claimed to have emerged through intensification and the arrival of entirely new issues in the future cannot be ruled out with a topic as transformative as digitalization.

*Moral Patients: Responsible towards Whom?*

Our second analytical question is about towards whom (or what) the firm is responsible. The established way of approaching this issue is to say that firms are responsible towards their stakeholders; groups and individuals who can affect the company or be affected by it (Freeman, 1984). Besides identifying stakeholders *per se*, an important consideration is stakeholder salience which determines the priority awarded to different stakeholder groups by managers. According to Mitchell, Agle and Wood (1997), stakeholder salience is based on the extent to which the stakeholders exhibit the characteristics of power, legitimacy, and urgency. The changes brought by digitalization can be examined in light of their influential framework.

First, the digital economy can *change the salience of existing stakeholder groups*. Power, legitimacy and urgency are all socially constructed concepts, and the model is dynamic, meaning that attributions of these characteristics can change with time (Mitchell et al., 1997). The advent of social media can be strategically used by stakeholders to influence these characteristics and thus build salience (Coombs & Holladay, 2015). For example, a stakeholder can make visible claims in social media and thus threaten the reputation of a firm (Ibid.) or amass the support of other stakeholders even from far-away contexts thanks to fast and cheap communications around the world. In an extreme case, a group could move from being a nonstakeholder with no salience (without any power, legitimacy, or urgency) to being a stakeholder. Note that the salience of a



stakeholder group could also decrease thanks to digital technology, for example if increased transparency takes away the legitimacy of its claims.

Second, digitalization may *make new stakeholder groups emerge*. This means that digitalization may invite us to question the definition of a stakeholder in a fundamental sense. The classic definition of stakeholders as ‘groups and individuals’ implicitly restricts stakeholders to human entities. According to a strict interpretation, even the natural environment is therefore not a stakeholder, but its interests need to be represented by others who are (though not all see it this way; see Haigh & Griffiths, 2009 for a discussion). However, with digital technology, the question arises whether algorithms and robots could be regarded as stakeholders. Gunkel (2007) notes that the scope of ethical consideration has only ever included humans, but as it has become more inclusive, potentially extending to animals, the next question is: What about the machine? Android ethics is already asking the question whether at some point machines are deserving of ethical consideration (Ibid.). Whether AI could be granted legal personhood was discussed as a theoretical question as early as 1992 by Solum, while Laukyte (2017) asks whether artificial agents should be granted rights similar to those accorded to group agents.

#### *Moral Agents: Who is Responsible?*

The simplest answer to our third analytical question would be that a company is responsible for its own direct actions. However, hand in hand with increased outsourcing and globalisation the scope of corporate responsibility has expanded to cover entire supply chains, and it is no longer sufficient for a firm to limit its responsibility considerations to its own actions only (see also Schrempf-Stirling, Palazzo, & Phillips, 2016). Companies are often held accountable for the labour practices of their suppliers in third countries. Macdonald (2011) discusses institutional relationships through

which firms may indirectly influence social outcomes: decentralised structures such as joint ventures, participation in business networks and supply chains, and relationships with governments and other actors with host countries. She concludes that “there remain important domains of business influence [...] for which the boundaries of business responsibility remain unclear” (p. 560).

The concept of ‘complicity’ has been brought forward to help identify corporate responsibilities beyond the direct actions of the firm itself. Complicity, which describes how firms may contribute to social impacts through their relationships, has been discussed in the context of human rights, but it has broader applicability. Complicity can be direct (as when a company knowingly assists in a harmful action), beneficial (when it benefits from the harmful action committed by someone else), or silent (when it fails to raise the issue with appropriate authorities) (ISO 26000, 6.3.5.1). Applying similar thinking in the context of online data supply chains, Martin (2016) argues that firms should consider their responsibilities in entire systems where they benefit from harmful practices, have the ability to stop harmful practices, or played a part in implementing harmful practices.

The digital economy can affect the allocation of responsibility among actors in at least two ways. First, it can *blur answers to the question of where responsibility resides between firms and individuals*. Already there is the longstanding (and unresolved) debate in business ethics on whether firms or only individuals can be assigned moral responsibilities (e.g., Orts & Smith, 2017). We can leave aside this debate, however, if we take the position that the entity of the firm can be conceived of as either the ethical agent in itself or through its individual employees as part of a collective. Less readily addressed are the implications of this question for where responsibility resides in multi-sided markets where individuals can participate in multiple roles beyond that of a

traditional consumer. For example, in the sharing economy, should the responsible actors be seen to be the firms (operating through platforms, with employees using those platforms) or the citizens (participating in peer-to-peer transactions—as potentially both providers and consumers—aided by and taking advantage of the platforms provided by firms)? Further, can firms such as Facebook and Twitter be considered only neutral intermediaries for communications, or do they have some responsibility for the content that third parties (including individuals) post on their platforms? What of the growing use of these platforms in human trafficking and ransom demands (Van Reisen & Mawere, 2017)? And what about the responsibility for mobile applications which are sold through Google Play or AppStore but whose development ecosystem may be quite deprofessionalised, including application developers who are hobbyists with little knowledge of privacy considerations (Shilton & Greene, 2019, p. 132)?

Second, the digital economy can *raise novel questions concerning where responsibility resides between humans and machines*. Will a firm that is producing or utilizing artificial agents be responsible for the decisions and actions of those agents? Even when the artificial agents are self-learning and autonomous, and the firm could not predict how they would behave in a particular situation? (Johnson, 2015; Matthias, 2004). And which firm would be held responsible, the developer or the user of the technology? Matthias (2004) argues that a ‘responsibility gap’ is created when moral responsibility for the actions of an autonomous system cannot be assigned to any humans. But should the machine itself be responsible – can moral responsibility be assigned to a machine? According to Hellström (2013), decisive for the assignment of moral responsibility to an artificial agent is the extent of its autonomous power, that is, “the amount and level of actions, interactions and decisions an agent is capable of performing on its own” (p. 101). Humans also have a tendency to assign more responsibility to actors that are human-like (ibid.).

### *Summary of Impacts on CR Domain*

Combining our key digital economy phenomena and the three core CR questions, Figure 2 summarizes the ways in which the digitalization of the economy may impact the domain of CR. However, the extent to which these impacts challenge current assumptions varies, as will be discussed below.

---

Insert Figure 2 about here

---

The bottom part of Figure 2 contains two horizontal boxes. One shows the management implications of the digital economy on the field of CR. There are immediate implications across the entire CR field, originating from all the digital economy phenomena. As we have discussed, and as the examples in the figure illustrate, there are changes relating to what a firm may be responsible for, towards whom it may be responsible, and who is responsible. Managers (and management researchers) need to be aware of these changes. The potential change identified in Figure 2 without immediate management implications relates to the question of having bots and robots as stakeholders, which is at the moment only speculative.

In this paper, however, our key interest is in potential *fundamental* changes to the field (see the other horizontal box in Figure 2). Even though there are clear changes throughout the CR field, a closer examination reveals that not all of them challenge fundamental assumptions in the current theoretical understanding of CR. As regards the first question about moral actions, CR issues may now be manifest in a different form, or possibly even be intensified, with a new urgency. Nonetheless, existing conceptions about what kinds of issues firms are responsible for can still accommodate the issues arising with the digital economy. The ‘new’ issues that we might

currently identify are ultimately intensified forms of ‘old’ issues. There is no requirement to re-examine foundational assumptions – unless entirely different, new issues emerge in the future.

As to the second question about moral patients, insofar as we are talking about changed salience of stakeholder groups, this is again fully catered for within extant assumptions. Stakeholder salience is known to be dynamic (Mitchell et al., 1997), and even if the changes in salience now occur because of digitalization, this does not challenge fundamental assumptions. If we are talking about awarding moral consideration to artificial agents, that would certainly call for a re-examination of the fundamental assumption concerning who can be a stakeholder (stakeholder theory currently being about human actors and their interactions; Freeman et al., 2020). For now, this possibility remains speculative. (Note that if machines can be moral agents under some circumstances, proponents of this view should logically ascribe them with the status of moral patients in certain circumstances too.)

In contrast to our first and second questions, the third question about moral agents (“Who is responsible?”) gives rise to the greatest need to re-examine fundamental assumptions, as we explain in the next section, ultimately concluding with a research agenda (Table 3).

## ALLOCATING RESPONSIBILITY AMONG MORAL AGENTS IN THE DIGITAL ECONOMY

We now examine the third question in greater depth and explore what the re-examination of those fundamental assumptions might mean for theory development and a future research agenda in the domain of CR in the digital economy and, more specifically, moral agency in a digital world.

The research agenda can be approached from two distinct perspectives, the normative and the descriptive. The dominant feature of a normative perspective on CR is “its emphasis on

formulating prescriptive moral judgments” (Trevino & Weaver, 1994, p. 114). A normative examination of allocating responsibility in the digital economy would be conceptual and effectively ask where the responsibility “correctly” lies—*who should be* responsible—generally with recourse to ethical theory. By contrast, the descriptive perspective aims to “answer questions about *what is* by attempting to describe, explain, and/or predict phenomena in the empirical world” (Trevino & Weaver, 1994, p. 117). A descriptive examination of allocating responsibility in the digital economy would ask where people perceive responsibilities to lie and generally rely on empirical observation.

Both perspectives are essential: the normative perspective helps managers and other actors to identify the morally right course of action and to determine their individual moral responsibility for outcomes; while the descriptive perspective helps those managers and other actors to understand the perceptions and judgments that stakeholders make about the actions undertaken and who is responsible for those actions. Accordingly, we develop the research agenda from both perspectives. This also highlights the important research question about whether there is any misalignment between normative and descriptive views and findings.

### *Re-examining the Assumption of Human-Only Actors*

What, then, are the key underlying assumptions that need to be re-examined in the digital economy, in relation to our third question? We identify two such assumptions. The first change brought by the digital economy is the need to *remove the assumption of human-only actors*. The assumption of human-only actors is well evident in the literature that examines corporate moral agency. Even though there is the debate about whether we can ascribe moral responsibility to organizations per se, or only to the individuals that make up the organization (Orts & Smith, 2017), proponents of

both views start from assumptions about *human* actions and intentions, be they as individuals or as a collective. However, allocating responsibility requires attention to the decisions and actions of *machines* in the digital economy.

Two fatal crashes of the Boeing 737 MAX airplane illustrate the limitations of the assumption of human-only actors (see also Lindebaum, Vesa, & den Hond, 2020). While the causes of the accidents are multiple and include failings by individuals (e.g., the plane design engineers; the pilots; the air traffic controllers) and individuals as part of a collective (e.g., lax maintenance culture and practices at Lion Air; the overreliance of regulators on assurances from the plane maker), there is also the role of the Maneuvering Characteristics Augmentation System (MCAS) which repeatedly forced down the nose of the plane after receiving faulty information from a sensor, despite the ultimately unsuccessful efforts of the pilots to override it, and it is at least the primary cause of both crashes (Langewiesche, 2019). MCAS is essentially a machine that incorporates AI. Similarly, the controversy over the Apple credit card that assigned lower credit limits to female applicants over male applicants, despite otherwise similar credit scores, apparently stemmed from machine learning and algorithms that generated the credit limits without human oversight (BBC News, 2019).

Hence, machines can clearly be actors with causal responsibility, but what does this mean for moral responsibility? It has been postulated that humans are always ultimately responsible (even if they are one or more steps removed and lack awareness of their responsibility), but also that humans cannot be responsible for actions by autonomous machines; that such machines might someday be morally responsible in a narrow sense; or that responsibility should be shared between machines and humans (Noorman & Johnson, 2014). However, while machines may appear to have ‘goals’ and in a sense they indeed do (a chess computer selects the moves to win), fundamentally

they do not have their own intentions and their goals are ultimately set by humans (the chess computer does not decide to play another game). This is an area with potential for discrepancy between normative ethical theory and layperson perceptions. Furthermore, questions of legal responsibility (which may not fully overlap with moral responsibility) are also important in shaping the workings of the digital economy. Whether machines are responsible, morally or legally, has implications for the extent to which humans and their organizations are responsible, and which humans or organizations. In sum, crucial questions about allocating corporate responsibility are opened up by the removal of the assumption of human-only actors.

It might be noted that machines have had potential causal responsibility before the arrival of the digital economy; an example would be an automated coffeemaker that can be pre-set to brew coffee in the morning, though it might also malfunction and set itself on fire. However, the CR literature did not give this much attention because of the limited scope of non-human actors: examples of machines having control were generally few and far between, the range of activities over which they had control was small, and human oversight was generally present or close at hand and with full understanding of what the machine was doing and how to intervene if needs be. This scope has changed out of all recognition with the machines operating in the digital economy.

### *Re-examining the Assumption of Clearly Defined Entities as Actors*

The second fundamental change brought by the digital economy is the need to *remove the assumption of clearly defined entities as actors*. With clearly defined entities we mean that an actor-entity is distinct, with clear organizational boundaries, that its role in the action is clear, and that its relationship with other actor-entities is clearly defined. Schrempf-Stirling et al. (2016, p.



702, emphasis added) capture the fact that such assumptions are indeed attached to corporate moral agency, summarizing:

Key characteristics of corporate moral agency are that the agent is able to act, has intentions, is able to reflect on those intentions (Werhane, 1985), has a moral relationship with others (Donaldson, 1982), *and is a single entity or unit to whom responsibilities can be assigned* [...] The corporation *is a distinct entity* and has moral agency, which means it is able to take responsibility for actions within its locus of control. In the debate on CSR (Matten & Moon, 2008), corporate citizenship (Moon et al., 2005), and stakeholder theory (Harrison, Bosse, Phillips, 2010), *scholars routinely take this assumption for granted* when discussing responsibilities of corporations.

Yet this characterization of business as involving clearly defined entities is becoming increasingly outdated. Jacobides, Cennamo and Gawer (2018, p. 2256) define an ecosystem as “a group of interacting firms that depend on each other’s activities.” This might sound familiar—Nespresso is an ecosystem (Jacobides, 2019)—but there has been a surge of interest in recent years. This is associated in large part with the growth of the digital economy. So Fuller, Jacobides and Reeves (2019) speak of an ecosystem as a substantive new phenomenon that has accompanied the emergence of the digital economy, referring to “the rise of dynamic, multicompany systems as a new way of organizing economic activity.” This is critical in thinking about allocating responsibility because “the firm is no longer an independent strategic actor” (Jacobides, 2019, p. 131). If it no longer acts independently and, instead, is acting in concert with multiple parties, then it becomes more difficult to say who is to blame when something goes wrong.

Thus, in the digital economy, the assumption of clearly defined entities to whom responsibility can be assigned becomes problematic for a number of reasons. First, organizational boundaries become permeable with new organizational forms such as platforms and ecosystems, including “independent contractors performing contingent, part-time, and temporary work” (Laamanen, Pfeffer, Rong, & Van de Ven, 2018, p. 216). Second, the roles of producers vs. consumers, or those of employees vs. contractors, become fuzzy in multi-sided markets. Third, outcomes tend to arise from interactions of disperse actions by multiple parties – potentially much more numerous than before because of how digitalization enables parties to come together in a different way – though not always with clear governance mechanisms (Jacobides et al., 2018). All this serves to obscure the identification of which actors there are with what boundaries that could be assigned responsibility, and what is their contribution and role in the matter and thus what share of responsibility could be apportioned to them. Again, therefore, important questions about allocating corporate responsibility are opened up by the removal of this second assumption.

#### *Towards a Research Agenda on Moral Agency in a Digital Economy*

To consider the individual and joint impacts of removing these two assumptions we can build a matrix as in Table 2. The table helps to identify actors that are involved in undertaking a moral action. Since our key question is “who is responsible?” we need to establish the nature of the entities taking the moral actions. This is so we can discuss whether these entities can be moral agents, to which ones among them moral responsibility can be assigned, and how it can be apportioned between them in case of shared responsibility. When both assumptions remain in place, the entity or entities involved are clear, so they can be identified, and they are all human, so there is no question about their moral agency (keeping in mind our earlier note about the moral

agency of corporations). Responsibility can be assigned to the actor or apportioned between the actors. This is the situation illustrated in quadrant I of Table 2.

---

Insert Table 2 about here

---

In quadrant II, there is certainty about the entities involved in the moral action, as in the pre-digital economy context, but now the entities (or at least some of them) are machines. This is where decision-making could be given over entirely to machines, with no effective human oversight. For example, Babic, Gerke, Evgeniou, and Cohen (2019, p. 1202) describe AI used for diagnosis in health contexts but presenting a dilemma to regulators: “should the regulator limit its authorization to market only the version of the algorithm that was submitted, or permit marketing of an algorithm that can learn and adapt to new conditions?” This is important because machine learning of, say, predictions of breast cancer, may improve as more data are analyzed. As Babic et al. (2019, pp. 1202-1203) note: “Parametric updates are at the core of modern AI/ML systems—they take place almost continuously, without human input, and their effects can be hard to identify ex-ante.”

In quadrant III, there is only human input without AI/machine decision-making, but multiple entities are involved and allocating responsibility between these entities is difficult because of their complex, interconnected roles within the ecosystem. The possibility of what Floridi (2016) describes as “distributed environments” is not new in itself. Multiple actors in apparel supply chains bear some responsibility for the Rana Plaza tragedy, where a building collapsed killing over 1100 workers. Even if much of the blame can be attributed to the building owner who added floors to the building without planning permission, other parties were involved with uncertain degrees of responsibility, such as factory owners who refused to let workers leave

the building when it started cracking (De Los Reyes, Scholz, & Smith, 2017). This scenario has become increasingly common with the digital economy. For example, while Airbnb might claim it is merely a platform connecting property owners with consumers looking for short stay accommodation, the question arises whether it bears responsibility for issues that develop as a consequence, such as racial discrimination against some would-be customers or bad behaviour by customers who, for example, host noisy gatherings in residential apartment buildings. In both examples there are multiple responsible parties, which becomes all the more likely in the digital economy.

Finally, in quadrant IV, there is the possibility of a combination of quadrants II and III, with multiple blurred entities involved, both human and machine. A simple illustration of this is the self-driving car scenario with human oversight. In March 2018, an autonomous car driving at 40 mph in Tempe, Arizona struck and killed a woman pushing her bicycle in the street (Wakabayashi, 2018). The self-driving vehicle was undergoing tests with a backup driver behind the wheel who had not intervened. Does responsibility lie with the backup driver, the developers of the technology, the city government that had approved the tests or, perhaps, with the vehicle itself? This quadrant also includes machine-machine interactions, as in the earlier Boeing 737 Max example (interaction of sensors and MCAS) or with the Internet of Things.

To sum up, the result of the new quadrants largely stemming from the growth of the digital economy (II to IV) is reduced clarity about how responsibility can be assigned to, and apportioned among, different human and non-human actors. Figure 3, which derives from Table 2, outlines candidates for allocating responsibility in different situations. The greyed area indicates what is new in the digital economy: anything that involves machine actors, but also new kinds of complexity arising from more intricate human entities, as discussed above.

---

Insert Figure 3 about here

---

Based on our examination of moral agency in the digital world, we can now identify a research agenda with regard to the greyed area in Figure 3, which represents the novel situations in a digital economy. On the basis of the foregoing analysis, Table 3 presents a research agenda, summarizing the research questions that need to be answered in order to be able to adequately address the question “Who is responsible?” in the digital economy.

---

Insert Table 3 about here

---

## DISCUSSION AND CONCLUSIONS

In this article, we set out to examine how the digital transformation in the economy may affect the field of CR and whether some of its foundational assumptions need to be re-examined. We did this by identifying five key digital economic phenomena from a CR-relevant perspective, outlining their associated ethical opportunities and threats, and mapping those opportunities and threats against three core questions that define the basic landscape of the CR field. The three questions are responsible for what, responsible towards whom, and who is responsible? We found that while there are foreseeable changes with regard to all three questions, the first question does not challenge any foundational assumptions; the second one might in principle but this is quite speculative; and, finally, that the third one does challenge some important foundational assumptions.

The foundational assumptions that are challenged in the digital economy are that of human-only actors, and that of clearly defined entities as actors. Not being able to rely on these

assumptions can – compounded with the other changes to the CR domain that we have identified – lead to obscured responsibilities. From a societal viewpoint, an obscuring of responsibilities is a problem if ethical lacunae arise for which it is not clear which actor should be held responsible, or indeed if no actor can be held responsible. From a business viewpoint, obscured responsibility is problematic if firms cannot be clear about where they may or may not be held responsible.

### *Theoretical Contribution*

Speaking about how to improve theory in corporate responsibility, Wang, Gibson, and Zander (2020, p. 3) write that “we need theoretical efforts that either provide precision by being narrowly focused with clearly defined constructs and assumptions [...] or offer an integration of various contingencies and perspectives, providing a comprehensive understanding [...] under one unifying framework.” It is this latter route that we have taken in our paper. An increasing number of focused studies are available on individual CR-relevant issues within the broad topic of digital economy, but what has been lacking is an examination of the entire CR field and the extent to which its basic foundations may be challenged by digitalization. Thus, by taking this less-travelled route we contribute to filling an important gap in the literature.

To our knowledge, this is the first study from the vantage point of holistically examining the field of CR through the lens of digitalization and from the *reverse* perspective of how it is impacted by digitalization (‘CR meets digitalization’), rather than from that of how digitalization is impacted by CR considerations (‘digitalization meets CR’). Through this novel viewpoint and a systematic approach, our article contributes to the advancement of the field of CR. Johnson (2015) argues that a new technology typically tests and modifies existing regulatory frameworks, and Gunkel (2007, p. 176), writes (see also Floridi & Sanders, 2002):

The machine question [...] puts in question the entire edifice of ethics [...] not only adds interesting new dimensions to old problems, but leads us to rethink, methodologically, the very grounds on which our ethical positions are based [...].

Similarly, exposing the edifice of CR to the challenge of digitalization serves as an “acid test” that can reveal eventual gaps or weaknesses. Our findings show that it is in the domain of the third question – who is responsible? – where the current CR edifice is least robust. To answer questions about how to allocate moral responsibility to firms in situations where some of the actors are machines and where causal responsibility is distributed and possibly unclear, demands recourse to moral philosophy (Floridi 2016). While it is beyond the scope of this article to actually solve these questions, our contribution is to show how the CR domain may change in the digital economy and which of the changes are fundamental enough to reveal where the CR edifice needs updating. In fact, better understanding of the question “who is responsible?” has broader significance in CR even if it is digitalization that brings it to greater prominence.

### *Implications for Research*

Our findings have direct implications for future CR scholarship. Research on CR needs to stay abreast of the changing landscape of CR, whether the changes can be accommodated within current assumptions or whether they are of a more fundamental nature. This necessitates drawing from both CR literature and digitalization literature for best results (see also Fisher, List, Slavkovik, & Winfield, 2016; Flyverbom et al., 2019; Rahwan et al., 2019). We provide an organizing framework (see Figure 2) which helps to delineate where these research topics fit within the field,

while Table 3 spells out a research agenda in relation to the final column of the framework, highlighting many directions for further research.

In particular, we have shown how allocating responsibility between different human (including organizational), and potentially also non-human, actors is not yet on firm ground. For example, various positions have been put forward concerning the responsibility of autonomous robots (Noorman & Johnson, 2014), and the concept of complicity has received more attention recently in the context of corporate responsibility (e.g., Monge, 2015). Further research – both descriptive and normative – is needed to address questions of allocating responsibility between various actors in dispersed or otherwise obscured situations, as we have outlined in detail in the research agenda we propose with regard to this question. These topics could be addressed conceptually but also empirically, attempting to tease out how people perceive the limits of responsibility, with experiments potentially providing interesting insights.

### *Implications for Managers*

General managers, managers tasked with digitalization, and CR managers can all benefit from our findings in helping them to think systematically about CR in a digitalized world. Digitalization will touch upon practically every firm: either because they employ digital technologies themselves or, if not, because they operate in an economy where most others do. To stay abreast of digital developments and understand what they mean for your company is a challenge for all, but it is perhaps less evident to CR managers that they also should fully understand digitalization and its impacts.

Two kinds of management implications arise from the intersection of digitalization and CR. First, implications may be given from the perspective of improving responsibility. To be able



to fully implement CR in a digitalizing economy, managers need to be digitally literate and able to understand the ethics-relevant changes brought by digital technology. In other words, CR managers need to understand how digitalization may change the CR issues that the firm needs to address (e.g., what are the issues, how do they manifest themselves?), the stakeholders towards whom the firm needs to be responsible (who or what are they, which ones are more important than others?), as well as who is responsible in various scenarios and where are thus the limits of the firm's responsibility. A deep understanding is required for two reasons: the changes brought by digitalization are not only technical and superficial but can disrupt business in fundamental ways; and the technologies are developing fast. The framework that our article offers can help practitioners to better navigate in the world of CR in a digital economy.

Second, there are managerial implications from a risk-management perspective. Managers need to understand the changes in responsibilities brought by digitalization so that they can negotiate appropriate contracts with their suppliers and clients. They also need to understand how developments in digital technology may result in outcomes that invite resentment or resistance based on ethical concerns. Unless sufficiently addressed through self-regulation and CR, these concerns may justifiably invite regulation, also affecting the ability of firms to do business. For instance, a perception of being under surveillance may harm employee morale; a lack of trust concerning the protection of personal data may make a customer turn to a competitor. Discussions have started of regulating digitalization (e.g., there are calls for independent audits of algorithms, or for regulations on labour conditions in the sharing economy). In sum, an understanding of how digitalization is shaping the field of CR is required to avoid harmful business implications.

Digital technology is developing rapidly, bringing major transformations in ways of doing business. To remain relevant, the field of CR must update itself consistent with these

developments, extending, as we have shown, to foundational assumptions in the field and with profound implications for researchers and practitioners.

## REFERENCES

- Andriole, S. J. 2017. *Five myths about digital transformation*. MIT Sloan Management Review, 58(3): 20-22.
- Arnaboldi, M., Busco, C., & Cuganesan, S. 2017. *Accounting, accountability, social media and big data: revolution or hype?* Accounting, Auditing & Accountability Journal, 30(4): 762-776.
- Austin, R. D., & Upton, D. M. 2016. *Leading in the age of super-transparency*. MIT Sloan Management Review, 57(2): 24-32.
- BBC News. 2019, November 11. *Apple's 'sexist' credit card investigated by US regulator*. BBC News. <https://www.bbc.com/news/business-50365609>.
- Babic, B., Gerke, S., Evgeniou, T., & Cohen, I. G. 2019. *Algorithms on regulatory lockdown in medicine*. Science, 366(6740): 1202-1204.
- Baesens, B., De Winne, S., & Sels, L. 2017. *Is your company ready for HR analytics?* MIT Sloan Management Review, 58(2): 20-21.
- Calo, R. 2014. *Digital market manipulation*. The George Washington Law Review, 82: 995-1051.
- Castelló, I., Etter, M., & Årup Nielsen, F. 2016. *Strategies of legitimacy through social media: The networked strategy*. Journal of Management Studies, 53(3): 402-432.
- Constantiou, I. D., & Kallinikos, J. 2015. *New games, new rules: big data and the changing context of strategy*. Journal of Information Technology, 30(1): 44-57.
- Coombs, W. T., & Holladay, S. J. 2015. How activists shape CSR: Insights from Internet contagion and contingency theories. In: *Corporate social responsibility in the digital age*. Developments in corporate governance and responsibility, 7: 85-97.
- De los Reyes, G., Scholz, M., & Smith, N. C. 2017. *Beyond the "win-win": Creating shared value requires ethical frameworks*. California Management Review, 59(2): 142-167.
- Donaldson, T., & Preston, L. 1995. *The stakeholder theory of the corporation: concepts, evidence, and implications*. Academy of Management Review, 20(1): 65-91.
- Etter, M., Colleoni, E., Illia, L., Meggiorin, K., & D'Eugenio, A. 2018. *Measuring organizational legitimacy in social media: Assessing citizens' judgments with sentiment analysis*. Business & Society, 57(1): 60-97.
- Etter, M., Fieseler, C., & Whelan, G. 2019. *Sharing economy, sharing responsibility? Corporate social responsibility in the digital age*. Journal of Business Ethics, 159(4): 935-942.

- Faraj, S., Pachidi, S., & Sayegh, K. 2018. *Working and organizing in the age of the learning algorithm*. Information and Organization, 28: 62-70.
- Fisher, M., List, C., Slavkovik, M., & Winfield, A. 2016. *Engineering moral machines*. Informatik Spektrum, 39(6): 467-489.
- Floridi, L. 2016. *Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions*. Philosophical Transactions of the Royal Society A, 374(2083): 1-13.
- Floridi, L., & Sanders, J. W. 2002. *Mapping the foundationalist debate in computer ethics*. Ethics and Information Technology, 4: 1-9.
- Floridi, L., & Sanders, J. W. 2004. *On the morality of artificial agents*. Minds and Machine, 14: 349-379.
- Flyverbom, M., Deibert, R., & Matten, D. 2019. *The governance of digital technology, big data, and the Internet: new roles and responsibilities for business*. Business & Society, 58(1): 3-19.
- Freeman, R. E. 1984. *Strategic management: A stakeholder approach*. Boston, MA: Harper Collins.
- Freeman, R. E., Harrison, J. S., Wicks, A. C., Parmar, B., & de Colle, S. 2010. *Stakeholder theory: The state of the art*. Cambridge, UK: Cambridge University Press.
- Freeman, R. E., Phillips, R., & Sisodia, R. 2020. *Tensions in stakeholder theory*. Business & Society, 59(2): 213-231.
- Fuller, J., Jacobides, M. G., & Reeves, M. 2019. *The myths and realities of business ecosystems*. MIT Sloan Management Review (digital edition), reprint 60315. See: <https://mitsmr.com/2Nt8fmC>.
- Gerwe, O., & Silva, R. 2020. *Clarifying the sharing economy: conceptualization, typology, antecedents, and effects*. Academy of Management Perspectives, 34(1): 65-96.
- Gray, K., & Wegner, D. M. 2009. *Moral typecasting: divergent perceptions of moral agents and moral patients*. Journal of Personality and Social Psychology, 96(3): 505–520.
- Gunkel, D. J. 2007. *Thinking otherwise: Ethics, technology and other subjects*. Ethics and Information Technology, 9: 165-177.
- Haigh, N., & Griffiths, A. 2009. *The natural environment as a primary stakeholder: The case of climate change*. Business Strategy and the Environment, 18(6): 347-359.

- Hellström, T. 2013. *On the moral responsibility of military robots*. Ethics and Information Technology, 15: 99-107.
- Illia, L., Romenti, S., Rodriguez-Cánovas, B., Murtarelli, G., & Carroll, C. E. 2017. *Exploring corporations' dialogue about CSR in the digital era*. Journal of Business Ethics, 146(1): 39-58.
- ISO. 2010. *Guidance on Social Responsibility – ISO 26000:2010*. Geneva: International Organization for Standardization.
- Jacobides, M.G. 2019. *In the ecosystem economy, what's your strategy?* Harvard Business Review, 97(5): 128-137.
- Jacobides, M.G., Cennamo, C., & Gawer, A. 2018. *Towards a theory of ecosystems*. Strategic Management Journal, 39: 2255-2276.
- Jersinovic, J. 2017, June 26. *The ethics of using AI in advertising*. AdAge. <http://adage.com/article/digitalnext/ethics-ai-advertising/309535/>.
- Johnson, D. G. 2015. *Technology with no human responsibility?* Journal of Business Ethics, 127: 707-715.
- Jones, R. 2017, July 24. *Roomba's next big step is selling maps of your home to the highest bidder*. Gizmodo. <http://gizmodo.com/roombas-next-big-step-is-selling-maps-of-your-home-to-t-1797187829>.
- Laamanen, T., Pfeffer, J., Rong, K., & Van de Ven, A. 2018. *Editors' introduction: business models, ecosystems, and society in the sharing economy*. Academy of Management Discoveries, 4(3): 213-219.
- Langewiesche, W. 2019, September 18. *What really brought down the Boeing 737 Max?* The New York Times Magazine. <https://www.nytimes.com/2019/09/18/magazine/boeing-737-max-crashes.html>.
- Laukyte, M. 2017. *Artificial agents among us: Should we recognize them as agents proper?* Ethics and Information Technology, 19: 1-17.
- Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. 2015. *Working with machines: The impact of algorithmic and data-driven management on human workers*. Conference on Human Factors in Computing Systems - Proceedings, April, pp. 1603-1612.
- Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. 2019. *The challenges of algorithm-based HR decision-making for personal integrity*. Journal of Business Ethics, 160(2): 377-392.

- Lindebaum, D., Vesa, M., & den Hond, F. 2020. *Insights from “The Machine Stops” to better understand rational assumptions in algorithmic decision making and its implications for organizations*. *Academy of Management Review*, 45(1): 247-263.
- Lyytinen, K., & Grover, V. 2017. *Management misinformation systems: A time to revisit?* *Journal of the Association for Information Systems*, 18(3): 206-230.
- Macdonald, K. 2011. *Re-thinking ‘spheres of responsibility’: business responsibility for indirect harm*. *Journal of Business Ethics*, 99: 549-563.
- Martin, K. E. 2015. *Ethical issues in the Big Data industry*. *MIS Quarterly Executive*, 14(2): 67-85.
- Martin, K. E. 2016. *Data aggregators, consumer data, and responsibility online: Who is tracking consumers online and should they stop?* *The Information Society*, 32(1): 51-63.
- Martin, K. E., & Freeman, R. E. 2003. *Some problems with employee monitoring*. *Journal of Business Ethics*, 43: 353-361.
- Matthias, A. 2004. *The responsibility gap: Ascribing responsibility for the actions of learning automata*. *Ethics and Information Technology*, 6: 175-183.
- Mitchell, R. K., Agle, B. R., & Wood, D. J. 1997. *Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts*. *Academy of Management Review*, 22: 853-886.
- Monge, R. 2015. *Institutionally driven moral conflicts and managerial action: dirty hands or permissible complicity?* *Journal of Business Ethics*, 129: 161-175.
- Noorman, M., & Johnson, D. G. 2014. *Negotiating autonomy and responsibility in military robots*. *Ethics and Information Technology*, 16: 51-62.
- O’Connor, S. 2016, September 8. *When your boss is an algorithm*. *Financial Times*.  
<https://www.ft.com/content/88fdc58e-754f-11e6-b60a-de4532d5ea35>.
- O’Reilly, T. 2016. *Managing the bots that are managing the business*. *MIT Sloan Management Review*, 58(1): 9-10.
- Orts, E. O., & Smith, N. C. (Eds.). 2017. *The moral responsibility of firms*. Oxford, UK: Oxford University Press.
- Parkes, D. C., & Wellman, M. P. 2015. *Economic reasoning and artificial intelligence*. *Science*, 349(6245): 267-272.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J., Breazeal, C., Crandall, J., Christakis, N., Couzin, I., Jackson, M., Jennings, N., Kamar, E., Kloumann, I.,

- Larochelle, H., Lazer, D., Mcelreath, R., Mislove, A., Parkes, D., Pentland, A., Roberts, M., Shariff, A., Tenenbaum, J., & Wellman, M. 2019. *Machine behaviour*. Nature, 568(7753): 477-486.
- Rosenblat, A., Levy, K. E. C., Barocas, S., & Hwang, T. 2017. *Discriminating tastes: Uber's customer ratings as vehicles for workplace discrimination*. Policy & Internet, 9(3): 256-279.
- Schildt, H. 2017. *Big data and organizational design – the brave new world of algorithmic management and computer augmented transparency*. Innovation, Organization & Management, 19(1): 23-30.
- Schrempf-Stirling, J., Palazzo, G., & Phillips, R. A. 2016. *Historic corporate social responsibility*. Academy of Management Review, 41(4): 700-719.
- Seele, P., Dierksmeier, C., Hofstetter, R., & Schultz, M. D. In press. *Mapping the ethicality of algorithmic pricing: A review of dynamic and personalized pricing*. Journal of Business Ethics. DOI: 10.1007/s10551-019-04371-w
- Shilton, K., & Greene, D. 2019. *Linking platforms, practices, and developer ethics: Levers for privacy discourse in mobile application development*. Journal of Business Ethics, 155(1): 131-146.
- Simpson, C. 2020, January 3. *Airbnb gained technology to judge 'narcissism' of guests*. The Telegraph. <https://www.telegraph.co.uk/news/2020/01/03/airbnb-gained-technology-judge-narcissism-guests/>.
- Singer, P. 1994. *Ethics*. Oxford: Oxford University Press.
- Solum, L. B. 1992. *Legal personhood for artificial intelligences*. North Carolina Law Review, 70(4): 1231-1287.
- Tapscott, D., & Ticoll, D. 2003. *The naked corporation*. New York: Free Press.
- Trevino, L., & Weaver, G. 1994. *Business ETHICS / BUSINESS ethics: one field or two?* Business Ethics Quarterly, 4(2): 113-128.
- Vaccaro, A., & Madsen, P. 2009. *Corporate dynamic transparency: The new ICT-driven ethics?* Ethics and Information Technology, 11: 113-122.
- Van Reisen, M., & Mawere, M. (Eds.) 2017. *Human trafficking and trauma in the digital era: the ongoing tragedy of the trade in refugees from Eritrea*. Cameroon: Langaa Research & Publishing.

- Wakabayashi, D. 2018, March 19. *Self-driving Uber car kills pedestrian in Arizona, where robots roam*. New York Times. <https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html>.
- Wang, H., Gibson, C., & Zander, U. 2020. *Editors' comments: Is research on corporate social responsibility undertheorized?* *Academy of Management Review*, 45(1): 1–6.
- West, S. M. 2019. *Data capitalism: Redefining the logics of surveillance and privacy*. *Business & Society*, 58(1): 20-41.
- Wood, D. J., Mitchell, R. K., Agle, B. R., & Bryan, L. M. In press. *Stakeholder identification and salience after 20 years: Progress, problems, and prospects*. *Business & Society*. DOI: 10.1177/0007650318816522



TABLE 1

Digital economy phenomena with examples of their associated ethical issues

Digital economy phenomenon (with examples of usage)	Examples of associated ethical issues	
	Opportunities	Threats
Digital marketing (Personalised targeting of messages; dynamic and personalized pricing)	Better understanding of consumer needs Lower prices for some consumers	Consumer privacy and autonomy Higher or unfair pricing for some consumers
Algorithmic management (Monitoring of employee or supplier performance; work scheduling)	Consistency in employee treatment	Employee freedom, privacy, respect; “Taylorism on steroids”
Autonomous processes in products and services (Algorithms (deciding e.g., on credit, recruitment); robots (e.g., drones, autonomous weapon systems, self-operating vehicles))	Machines acting without human weaknesses (such as cognitive limits, lack of self-control)	Machines different from humans (morals); similar to humans (bias); replacing humans (job loss, singularity); lack of explainability
Sharing economy (Uber, Lyft (transportation); Airbnb (accommodation); TaskRabbit (tasks))	Saving resources through sharing	Accountability; assignment of responsibility; bias in ratings; working conditions
Transparency and stakeholder governance (Availability and communication of information; dialogue and engagement)	(Super)transparency; traceability; accountability	Spreading false information; pressurizing companies in undesirable directions

TABLE 2

Nature of entities taking moral actions, after removing certain assumptions

		Remove assumption of human-only actors	
		No	Yes
Remove assumption of clearly defined entities as actors	No	I	II
		Clear human entities involved in the moral action	Clear human-machine or machine- machine entities involved in the moral action
	Yes	III	IV
		Blurred human entities involved in the moral action	Blurred human-machine or machine-machine entities involved in the moral action

TABLE 3

Research agenda for allocating responsibility in the digital economy

	Normative research questions	Descriptive research questions
Clear machine actors	<ul style="list-style-type: none"> <li>• Can machines have moral agency? Can they be responsible? Under what circumstances can they be responsible? In what sense be responsible? What kinds of machines can, if any? In what kind of contexts?</li> <li>• If machines can have ethics, whose / what kind of ethics should they be programmed, taught, trained to have? How can ethics be codified for machine use? Could machines code themselves? Would any of this be desirable?</li> <li>• What does it mean for a machine to be responsible - what does praise or blame mean for machines?</li> <li>• If machines can be moral actors, what implications does that have for them also being moral patients?</li> </ul>	<ul style="list-style-type: none"> <li>• How do people tend to perceive machines, as responsible agents or not? Are these perceptions compatible with the normative view?</li> <li>• How do perceptions of machine responsibility change between different groups of people / different machines / different contexts? Why?</li> <li>• How are perceptions of machine responsibility influenced when machines are anthropomorphized and personified?</li> <li>• How does the use of language affect people’s perceptions?</li> <li>• How are perceptions influenced by how the matter is being framed, and who might benefit from which kind of framings?</li> <li>• Do some actors deliberately try to “hide behind machines”?</li> </ul>
Blurred human-human actors	<ul style="list-style-type: none"> <li>• What should be the criteria for allocating responsibility between actors; e.g., in platform situations?</li> </ul>	<ul style="list-style-type: none"> <li>• How do people tend to perceive the responsibilities of various actors; e.g., in platform situations?</li> </ul>

Blurred human-machine actors	<p>All of the “clear machine actors” questions above, and in addition:</p> <ul style="list-style-type: none"> <li>• How to divide responsibility between the various humans and non-humans in the ecosystem: the developer, data provider, user, the algorithm or robot itself?</li> <li>• What if the actions of the machine are unpredictable or unexplainable (e.g., as a consequence of machine learning), how does that change the responsibility of the humans involved?</li> </ul>	<p>All of the “clear machine actors” questions above, and in addition:</p> <ul style="list-style-type: none"> <li>• How do people tend to perceive the responsibilities of humans vs. machines in blurred ecosystems constituted of various actors?</li> </ul>
Blurred machine-machine actors	<p>All of the “clear machine actors” questions above, and in addition:</p> <ul style="list-style-type: none"> <li>• When the machines themselves are multiple and integrated, how can responsibility be identified there?</li> </ul>	<p>All of the “clear machine actors” questions above, and in addition:</p> <ul style="list-style-type: none"> <li>• How do people tend to perceive the responsibilities of individual machines in integrated systems?</li> </ul>

FIGURE 1

Digital economy phenomena from a CR-relevant perspective

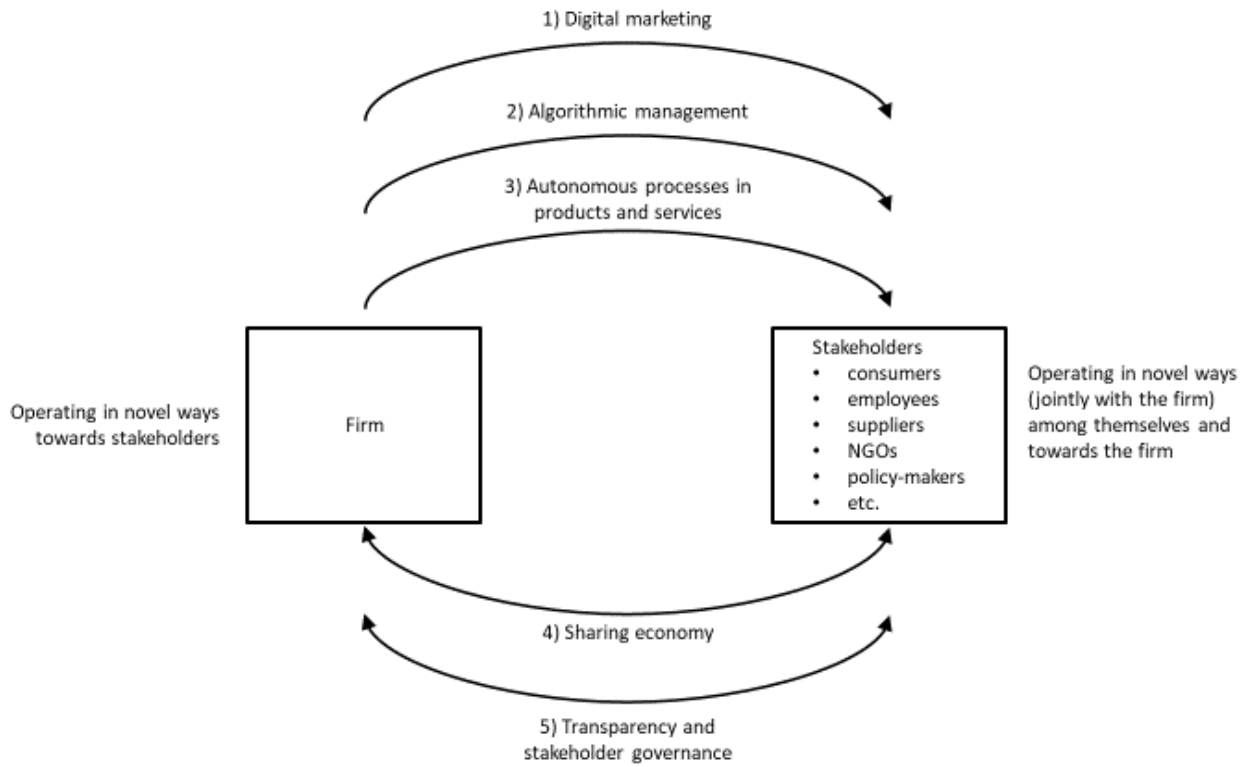


FIGURE 2

Illustrating how the digital economy can impact the field of CR

		Fundamental questions defining the landscape of CR			
		Moral actions: Responsible for what?	Moral patients: Responsible towards whom?	Moral agents: Who is responsible?	
Digital economy phenomena from a CR-relevant perspective (*)	Digital marketing	Issue manifesting in a novel way: consumer privacy	Issue intensified: consumer autonomy		
	Algorithmic management		Issue intensified: employee treatment		
	Autonomous processes in products and services	Issue manifesting in a novel way: job replacement	Issue alleviated or solved: AI and grand challenges	Potential new stakeholder groups: bots, robots	New considerations of responsibility between humans and machines
	Sharing economy	Issue manifesting in a novel way: working conditions in 'gig economy'			Blurring of responsibility between firms and individuals: multi-sided markets, traditional roles blurring
	Transparency and stakeholder governance	Issue manifesting in a novel way: transparency	Issue alleviated or solved: transparency	Changing salience of existing stakeholders: e.g. through social media	
<b>Management implications</b>		Immediate	Partly immediate, partly speculative	Immediate	
<b>Need to re-examine foundational assumptions</b>		None	Potential	Strong	

(\*) Note: Because the digital economy phenomena are mutually non-exclusive, their impacts could also appear in rows other than those indicated. For example, issues arising in the context of autonomous processes are also relevant, for example, to digital marketing and algorithmic management, as a result of autonomous processes being employed there.

FIGURE 3

Candidates for allocating responsibility, with the coloured area indicating what is new in the digital economy

