Research Article

URF PUBLISHERS
connect with research world

Journal of Artificial Intelligence, Machine Learning and Data Science

https://urfpublishers.com/journal/artificial-intelligence

Vol: 1 & Iss: 2

Artificial Intelligence and Corporate Digital Responsibility

Peter Jones[1] and Martin Wynn[2*]

[1]The Business School, University of Gloucestershire, Cheltenham, UK

[*2]School of Computing and Engineering, University of Gloucestershire, Cheltenham, UK

***Corresponding author:** Dr. Martin Wynn, School of Computing and Engineering, University of Gloucestershire, Cheltenham, UK. Email: mwynn@glos.ac.uk

Citation: Jones, P., & Wynn, M. (2023). Artificial Intelligence and Corporate Digital Responsibility. *J Artif Intell Mach Learn & Data Sci*, 1(2), 50-58.

Received: 12 March, 2023; **Accepted:** 17 April, 2023; **Published:** 20 April, 2023

A B S T R A C T

The continuing development of digital technologies is opening up a new range of opportunities, yet companies face, and may have to address, several new sets of responsibilities. However, to date, few companies have publicly reported on their digital responsibilities, and there has been little attention given to this subject in the academic literature. Adopting a qualitative inductive research approach, based on a scoping review of relevant literature and internet sources, and four case examples of major Information and Communication Technology (ICT) companies, this paper addresses two research questions. First, how do the ICT companies reconcile the evident benefits of Artificial Intelligence (AI), alongside the perceived challenges and concerns? Second, how are the ICT companies approaching their responsibilities regarding the development and deployment of AI products? The findings suggest that, although these companies are looking to acknowledge their social and technological responsibilities associated with AI, scant attention is paid to a number of significant issues relating to the responsible corporate management of AI. The paper makes a small contribution towards filling the gap in the literature on how companies are addressing Corporate Digital Responsibility (CDR) as regards AI product development and deployment, and can also be seen to contribute to the embryonic attempts to conceptualise CDR.

Keywords: Digital Technologies, Artificial Intelligence, AI, Corporate Digital Responsibility, CDR, ICT Companies, Environmental Responsibilities, Power Asymmetry

Introduction

Digital transformation – essentially the transitioning of business to new processes and business models via the adoption of digital technologies – is disrupting and reshaping traditional business operations and practices across the globe. However, while the continuing development of the digital technologies is opening up a new range of opportunities, companies face, and may have to address, several new sets of responsibilities. On the one hand, the digital technologies are seen to allow companies to accelerate efficient workflows, to increase profitability, and to deliver new services and value to their customers and clients. On the other hand, the corporate adoption of digital technologies also raises a wide variety social, economic, technical, and environmental concerns. Such concerns are increasingly being captured in the concept of corporate digital responsibility (CDR), simply defined simply defined as "a set of practices and behaviours that help an organisation use data and digital technologies in ways that are perceived as socially, economically, and environmentally responsible" (Corporate Digital Responsibility, 2022, para. 2).

There is growing public awareness of, and increasing concern about, the impact of Artificial Intelligence (AI). Elliot *et al*. (2021, p. 179), for example, suggested that "such concerns are revealed in public perceptions and uncertainty surrounding AI's future in society from technology executives overseeing the development and implementation of AI to the general public". More specifically, the call, supported by Elon Musk, one of the founders of the Open AI research laboratory, in March 2023 for a six-month moratorium on research into AI, to give time for widespread reflection, received massive public attention.

Indeed, Raja Chatila, professor emeritus at the Sorbonne and member of the working group of the French national digital ethics pilot committee, observed "everything that is currently happening in AI is taking place with no real ethical or legal controls. Companies are deploying tools on the web that have harmful effects" (Piquard, 2023, para. 8). To date, few companies have publicly reported on how they are discharging their digital responsibilities. This is not a problem *per se*, in that reporting on how companies are addressing such responsibilities is a voluntary process, and not subject to statutory regulation at present. At the same time, how companies are approaching and addressing CDR has received only limited attention in the academic literature.

In many ways AI is the most powerful of the digital technologies. Jyoti (2022, para. 1), for example, argued "AI is the most disruptive technology innovation of our lifetime", and that "enterprises are embracing AI and leveraging a variety of data types in all lines of business and industries", not least because "artificial intelligence forms the basis for all computer learning and is the future of all complex decision making" (NetApp, 2023, para. 1). Enholm *et al*. (2022) recognised AI as a wide-ranging set of technologies, and they suggested that there were two main ways of defining AI. The first "defines AI as a tool that solves a specific task that could be impossible or very time consuming for a human to complete", while the second "regards AI as a system that mimics human intelligence and cognitive processes, such as, interpreting, making inferences and learning" (p. 1712). In some ways this mirrors the distinction made between narrow AI and generative AI. The former is developed for a highly focused set of tasks, while the latter can create highly realistic and complex content. At the same time, within the business environment, there are three spheres within which AI is deployed. Firstly, they may be deploying AI within their own company, for example, to support advanced analytics to provide new customer insights, or to develop robots on production lines or digital twin applications in engineering or manufacturing. Secondly, they may be using AI outside of their company boundaries with customers, suppliers, business partners and the public at large, in applying algorithms to data made available via social media, e-mails and surveys, for example, to provide insights, projections and analysis of value to them. Thirdly, they may be developing and selling AI products. In all three contexts, company responsibilities for AI can be seen to represent a subset of CDR, and the authors sought to explore this dynamic in large ICT enterprises where all three environments are in evidence.

This exploratory paper draws material from four major Information and Communication Technology (ICT) multi-national companies to develop a narrative on how these large enterprises are addressing the responsibilities regarding the use and/or development of AI products, and as such, it offers a small contribution towards helping to fill the gap in the literature on how companies are addressing CDR. Following this introduction, the paper includes a literature review, a section on the research methodology, followed by an examination of the major themes identified in the four case examples that address the two main research questions. There then follows a discussion section and a conclusion.

Literature Review

CDR is emerging as a new theme within the academic literature, and while there have been calls for organisations to develop a better understanding of how to act in a digitally responsible manner, to date little work has been published on how individual companies are addressing CDR. Lobschat *et al*. (2021, p. 876), for example, argued "in the face of ethical challenges arising from the development and deployment of technology and data, organizations need to develop a better understanding of how to manage ethical dilemmas and overall act digitally responsible", and that CDR "requires tech companies, individual developers and designers, and any corporate actor employing digital technologies or data processing to be aware that the code they produce or deploy, as well as the data they collect and process, inherently create an ethical responsibility for them".

A number of definitions of CDR can be identified in the emerging literature. Lobschat *et al.* (2021, p. 875), for example, defined CDR as "the set of shared values and norms guiding an organization's operations with respect to the creation and operation of digital technology and data". For Elliot *et al.* (2021, p.184), CDR is "a voluntary commitment by organisations fulfilling the corporate rationaliser's role in representing community interests to inform good digital corporate actions and digital sustainability via collaborative guidance on addressing social, economic, and ecological impacts on digital society". From a more commercial perspective, Dentons (2022), the world's largest multi-national law firm, asserted that "CDR is designed to minimize the negative and maximize the positive impacts digitalization and digital tools can have on people and the environment". More succinctly, PWC (2023, para. 9) argued that it uses CDR "to design digitization in a responsible manner".

While research on CDR is still in its infancy in the academic literature, three sets of interlinked research themes, relevant to this paper on CDR and AI, can be identified within that literature: the principal area concerns responsibilities associated with the development of AI; then, there has been some focus on CDR in the service sector; finally attempts to conceptualise CDR also feature in the extant literature. The aim here is not to offer a comprehensive review of each of these themes, but rather to give some illustrative flavour of this research in order to provide a frame of reference and an academic context for the two research questions addressed in the case examples.

Firstly, there is growing research interest in the responsibilities associated with the development of AI. In looking to understand responsibility in AI, Constantinescu *et al*. (2021) highlighted the difficulties in assigning responsibilities to either technologies themselves or to their developers. Rather, responsible AI is generally taken to embrace a number of issues, ranging from the capabilities of the people involved in AI deployment to assuming a prospective ethical role, while embracing topics such as transparency, fairness, non-malignancy, and privacy. Constantinescu *et al*. (2021), however, claimed that within the context of AI ethics, the notion of responsibility does not automatically entail a moral dimension, and suggested that more attention must be paid to the responsibility of humans involved in AI design, and that designers need to become more context sensitivity to the moral demands of various communities. Such a prescription would also include ethical education which could enable designers to develop their moral imagination and critical thinking abilities, to help them make robust moral decisions about the development of AI.

Borenstein and Howard (2021) argued that the complex concerns emerging from the design and use of AI should serve as a reminder to focus on what developers and designers are learning about AI. Here, the argument is that it is important to train future members of the AI community, and other stakeholders, to reflect

on the ways AI might impact on people's lives and to embrace their responsibilities to enhance its benefits, while also mitigating its potential harms. By way of addressing future education and more specifically of nurturing a professional mindset in those who make up the AI community, Borenstein and Howard (2021) proposed three elements that could help to familiarise students with the emerging ethical challenges of AI: teaching the ethical design of algorithms; incorporating fundamental concepts of data science and the ethics of data acquisition, using real world data sets that require students to address privacy, fairness and legal issues while developing AI solutions; and, finally, regularly reinforcing the significance of ethics across the curriculum.

Turning from individual to political responsibilities, Djeffal *et al*. (2022) recognised that technologies based on AI posed a crucial policy challenge for policy makers, and looked to focus on how governments within the European Union looked to navigate and reconcile the potentially competing challenges of governing AI to stimulate innovation, mitigate risks and assume responsibility for the social impacts of AI. Their findings revealed multiple modes of governance within the European Union, with the major difference being between countries which follow a self-regulation approach, and those countries which combine entrepreneurial and regulatory approaches. In conclusion, Djeffal *et al*. (2022) argued that this finding qualifies dominant narratives in public and political debates, which suggest that policy makers have to choose between the promotion of emerging technologies and the regulation of the potential risks.

Wirtz *et al.* (2022), warned that the application of AI can pose serious risks to service customers as a result of ethical trade-offs made by companies in its design and operation. Here the authors argued "the use of AI in data driven business models that operate without oversight or accountability while covertly harvesting consumer data en masse……...could prompt addictive or coercive service interactions" (Wirtz *et al.* 2022, p.1). By way of a conclusion Wirtz *et al.* (2022, p. 4) claimed that "as the ethical risks, unintended consequences and trade-offs involved with the operation of increasingly powerful AI systems will escalate in future, service companies should ensure that robust and effective CDR practices are in place to manage these risks".

-Haenlein and Kaplan (2019, p. 5) defined AI as "a system's capacity to accurately read external input, learn from it, and use what it has learned to achieve specified objectives and tasks through flexible adaptation". Elliott *et al.* (2021), however, recognised that, while the increasing use of AI could improve productivity and efficiency, it inevitably brought costs involved in "delegating power to algorithmically based systems, some of whose workings are opaque and unobservable" (p.179). The authors looked to illustrate the opportunities and the threats of AI, *"*while raising awareness of Corporate Digital Responsibility (CDR) as a potential collaborative mechanism to demystify governance complexity and to establish an equitable digital society". Looking positively to the future, Elliott *et al.* (2021) argued that CDR "can potentially differentiate organisations, facilitating the gaining and maintaining of stakeholder trust and driving competitive advantage" (p.179).

Bonson *et al.* (2023) acknowledged that, as companies have looked to develop and apply AI to improve efficiency, reduce costs and personalise goods and services, and more particularly they have deployed disruptive AI tools with high levels of automation, the massive collection and processing of data and their inherent biases have raised concerns about human rights, data security and both privacy and ethical issues, across society. In the belief that there is an urgent need for more transparency in the use of AI, Bonson *et al.* (2023) analysed the current AI disclosure of 337 major Western European companies as outlined in their annual or sustainability reports. The authors findings revealed that AI disclosure was rather scarce and still at the preliminary stage, and that only 22 companies disclosed information related to automated decision making in their annual/sustainability reports.

In sharp contrast, Tigard (2021) challenged what he saw as the prevailing assumption that developments in AI systems are creating an ever-widening responsibility gap, which threatens to undermine both the moral and legal framework of society. Rather, Tigard (2021, p. 589) argued that there was "no techno-responsibility gap", and looked to demonstrate that moral responsibility was a dynamic and flexible process, that could effectively encompass emerging technological entities. Perhaps tangentially, Danaher (2022) argued that techno-responsibility gaps are sometimes to be welcomed, in that they make it possible to embrace such gaps. The gist of Danaher's (2022) argument is that human life is replete with morally tragic choices, that such choices are endemic to human decision making, that one potential advantage of AI is that tragic choice can be addressed directly, and that "we should favour delegation to autonomous systems without, necessarily, striving to fill the responsibility gaps that may be opened up as a result of that delegation" (p.26).

Cheng *et al.* (2022, para. 4) argued that "the dark sides of AI are receiving relatively little attention, especially from the academic community". In looking to take up this challenge, Mikalef *et al.* (2022) recognised that many large organisations have proposed a set of guiding principles around AI, but these principles have largely emerged from prominent cases where negative effects have been noted. Further, Mikalef *et al.* (2022) challenged some of the key assumptions around AI by utilising a dark side lens as a means of questioning the dominant paradigm around AI and suggesting a series of research questions. Here, the authors outlined, and updated, three sets of dark side assumptions and then drew out some emerging research questions. So, having updated "the dark side is known", to "the dark side is unknown", the authors (Mikalef *et al.,* 2022, Table 3) posed a number of "emerging research questions", including: "how can the fairness, transparency, accountability, and other features of responsible AI be measured when the issue in question is not known?", and "how can AI models be developed and trained to identify the unknown or accommodate issues that cannot be identified?"

Secondly, in the context of the services industries, Wirtz *et al*. (2021) acknowledged that the adoption of digital technologies carried serious privacy and ethical risks, and claimed that while CDR had been developed at the firm and society level, it had not been explored in a service context. Further, Wirtz *et al*. (2021) argued that the underlying causes of CDR issues result from the trade-offs between good CDR practices and corporate profit opportunities. They emphasised that service companies should look to ensure that CDR issues were addressed, particularly in supply chains, with their business partners, and where secondary users had access to their customers' data.

In reviewing CDR in service firms more generally, Wirtz *et al.* (2022a) looked to make a number of contributions to the literature. They demonstrated that CDR is critical in service contexts because of the enormous volume of data involved and the digital service technology's omnipresence and complexity; they synthesised the ethics and privacy literature using the CDR

data and technology life-cycle perspective to better understand the nature of these risks in a service context; to provide insights on the origins of these risks, by examining the digital service ecosystem and the related flows of money, service, data, insights, and technologies; and they proposed a set of strategies, tools and practices that service firms can use to build a strong CDR culture.

Thirdly, as regards the conceptualisation of CDR, Mihale-Wilson *et al.* (2022, p.1) argued that this was still in its infancy, and that their work sought "to contribute to CDR theory by providing a more in-depth conceptualization of the concept". Here, the authors theorized the link between the proposed CDR norms and digitization challenges, and argued these norms could serve as a preliminary conceptualisation of CDR. The eight norms included access; information and transparency; economic interests; and privacy and data security; and each norm had a descriptor. Thus, access relates to consumers having access to basic digital goods and services; in describing information and transparency, the focus is on consumers being able to locate appropriate information so they can be informed according to their individual wishes and needs; economic interests are described as the protection and promotion of the consumers' economic interests; while privacy and data security concerns the protection of consumers' privacy and the free flow of information, as well as the offer of protected and secure payment mechanisms.

Arguably more substantively, Bonson *et al.* (2023) identified, and summarised, three theoretical approaches that had been used to conceptualise CDR, namely voluntary disclosure theory, signalling theory, and legitimacy theory. Voluntary disclosure theory suggests that companies voluntarily disclose largely favourable information about their operations, and avoid the disclosure of unfavourable information. Here, the theory posits, the aims are to overcome a minimum mandatory level of disclosure, and to focus on favourable information to obtain positive economic impacts and to appeal to investors. Three main categories - intent, camouflage and need - are seen to be central to signalling theory. Intent is used to indicate a company's future plans in the hope of attracting investors; camouflage signals being employed to divert stakeholder attention away from a company's potential weaknesses; and need is used to influence potential investors' resource allocation decisions. Legitimacy theory is based around the belief that companies have to legitimate their practices in the eyes of society, and more specifically to ensure that they comply with stakeholder expectations.

At the same time, ethical theories also seem to be important in illuminating approaches to AI and CDR. Zoshak and Dew (2021), for example, suggested that two ethical theories, namely, deontology and consequentialism, dominated the testing and development of AI. Deontology takes a rule-based approach, and is based on a series of universal moral laws, while consequentialism is focused on the outcomes of behaviour. However, Zoshak and Dew (2021) argued that focusing on deontology and consequentialism in the development of AI systems risks encoding Western ethical reasoning into such systems at the expense of other forms of moral reasoning.

Given the issues raised in the above review of the existing literature concerning CDR and AI, this article addresses the following research questions (RQs) as they apply to four case examples of major ICT companies:

RQ1: How do the ICT companies reconcile the evident benefits of AI alongside the perceived challenges and concerns?

RQ2: How are the ICT companies approaching their responsibilities regarding the development and deployment of AI products?

Research Method

The research method was based on an initial scoping review of existing literature, followed by an assessment of four case examples of how major ICT companies were approaching their corporate responsibilities regarding AI. The scoping review method was first set out by Arksey and O'Malley (2005), and can be viewed as "a literature mapping process that allows a researcher to examine the 'landscape' of the literature based on a particular question of interest" (Hanneke *et al*. 2017, p.3). Armstrong *et al*. (2011) point out that there may be several different objectives in conducting such a review. These include an exploration of the extent of the literature, establishing the boundaries of the review, and identification of gaps in the extant literature. Scoping reviews can be used to lay the groundwork for subsequent research, and "are best employed when there is limited literature to inform the research question of interest" (Hanneke *et al*. 2017, p.5).

The research method was qualitative, and adopted an interpretivist paradigm, being based on an analysis of existing published literature and web sources. First, the authors reviewed recently published academic literature and information obtained from various research articles on Google Scholar, Scopus, Science Direct, and Web of Science to provide the material for the scoping review presented above. More than 80 relevant sources were located, which were studied and analysed manually to identify emerging themes and develop the two research questions. This was an iterative, cyclical process involving identification of the emergent themes and "reflective memoing and diagramming to ensure valid integration, interpretation, and synthesis of findings" (Finfgeld-Connett, 2014, p. 341).

Second, to address the research questions, the authors drew exclusively on the material on the responsible deployment of AI posted by the four case example companies - Google, Microsoft, Accenture, and IBM. These companies were identified in the scoping review as enterprises where AI products were developed or distributed, and where AI was used within the enterprise and with the customer base, thus covering all three environments noted in the Introduction section. All the cited material is in the public domain, and the authors judged it not necessary to seek permission to use appropriate quotations from these sources. Google had posted two sets of material, the one "Artificial Intelligence at Google: Our Principles", (Google, undated a) and the other entitled "Responsible AI Practices" (Google, undated b), written largely for AI systems designers. Microsoft posted a range of material on their approach to responsible AI, but the narrative below draws on two postings, namely "What is Microsoft's Approach to AI" (Microsoft, 2023 a) and "Responsible AI" (Microsoft, 2023b). Accenture posted "Artificial Intelligence" (Accenture, 2023) and IBM posted "AI Ethics in action: An enterprise guide to progressing trustworthy AI" (IBM, 2022) and "AI Ethics" (IBM, undated). The material from the companies does not always address the same issues, though there is some common ground in the narrative presented below. Rather, the narrative offers some insights into the four companies' public approaches to the responsibilities generated by their deployment of AI.

Google is a US multinational company focused on AI, cloud computing, computer hardware, computer software, quantum

computing, online advertising, e-commerce, search engine technology and the Internet. The company was founded in 1998 and it has some 140,000 employees. Microsoft is a US technology company and its product/service range includes computer hardware, software development, consumer electronics, cloud computing, video games, and social networking services. The company was founded in 1975 and has some 220,000 employees. The company's universally known software products are its Windows, Microsoft Office Suite, and the Internet Explorer and Edge web browsers. Accenture is an Irish-US professional services company specialising in ICT services and consulting. The company, originally founded in 1989 as Anderson Consulting, has a global presence and some 730,000 employees. IBM is a US based technology company, and its product/service range includes robotics, AI, cloud computing, blockchain, computer hardware, software, quantum computing and business consulting. The company was founded in 1911 and has some 280,000 employees.

CDR is a new theme in the business world and as noted in the Introduction to this paper, few companies have publicly reported on how they are approaching their digital responsibilities. However, the authors ascertained that four of the world's major ICT companies, notably, Google, Microsoft, Accenture and IBM, had posted some details of how they were addressing their responsibilities to AI on the Internet. As such, this paper might best be seen as an opportunistic endeavour designed to shed some preliminary light on an issue that has received very little attention in the academic literature. Each of the selected companies articulated their AI responsibilities in a variety of ways but rather than describing these commitments, the aim here is draw out, and illustrate, four sets of themes, which characterise how the companies are publicly addressing their AI responsibilities, as a narrative. It is therefore not intended that these cases should provide a basis for generalisation beyond the immediate company environment. As Flyvbjerg (2006) noted, each case should focus on the generation of a deep understanding of the complexity of the case, producing "concrete, context-dependent knowledge" (p. 223), rather than sweeping generalisations about the interplay of AI and CDR.

Findings

Building upon themes highlighted in the literature review and an analysis of the four companies' available documentation on AI and corporate responsibility, the two research questions are addressed below.

RQ1: How do the ICT companies reconcile the evident benefits of AI alongside the perceived challenges and concerns?

All four companies explicitly emphasised the benefits of AI. Under the banner "Responsibilities", Google (undated a, para.1), for example, claimed that the company "aspires to create technologies that solve important problems and help people in their daily lives. We are optimistic about the incredible potential for AI and other advanced technologies to empower people, widely benefit current and future generations, and work for the common NOT good." In a similar vein, Microsoft (2023a) asserted its beliefs that "AI is the defining technology of our time" (para. 1), and that "our AI tools and technologies are designed to benefit everyone at every level in every organisation" (para. 3). For Accenture (2023, para. 22) "AI helps businesses adapt at speed, with a regular stream of insights to drive innovation and customer advantage", as well as offering "improved accuracy and decision making", "empowered employees", and "superior customer service" (Accenture, 2023, para. 23).

This outward support for the benefits of AI was balanced by an evident awareness of the challenges and potential concerns raised by the growing deployment of AI, which was a common element in the selected companies' approach to their AI responsibilities. Accenture (2023, para. 25) claimed "AI is moving at a blistering pace and, as with any powerful technology, organisations need to build trust with the public and be accountable to their customers and employees". IBM (2022) argued that "business leaders need only scan the headlines to find examples of companies confronted with various societal, environmental, and political issues. Customers, employees, and even shareholders are more frequently demanding that organizations not only take a principled stance on current concerns, but also follow through with meaningful actions that lead to clear outcomes" (para. 4). Google (undated a, para. 3) recognised that AI technologies "raise important challenges that we need to address clearly, thoughtfully, and affirmatively".

RQ2: How are the ICT companies approaching their responsibilities regarding the development and deployment of AI products?

The four selected companies claimed that a number of principles guided their development of their responsibilities towards AI. These included: data privacy and security; fairness and inclusion; interpretability; accountability; safety; the avoidance of unfair bias; explainability; reliability; trust; and high standards of scientific excellence and control. While there is a good deal of commonality in these principles between the companies, similar principles were sometimes named, and expressed, differently. At the same time there was an awareness that the increasingly sophisticated development of AI and its diverse deployment in the business and social world demanded flexibility in responsibility policies. Google (undated a, para. 12), for example, acknowledged that the future development of AI "is dynamic and evolving, and we will approach our work with humility, a commitment to internal and external engagement, and a willingness to adapt our approach as we learn over time". Microsoft warned that as an AI system's performance can degrade over time, so it was important to measure its performance over time and retrain it as necessary.

In addressing fairness, Google claimed that fairness and inclusion were active and continuing elements in their AI development processes, and argued that while AI systems have the potential to be fairer and more inclusive, any unfairness in such systems can have wide scale impacts. Further, the company recognised that developing AI systems that are fair and inclusive for all is a hard task, and that if AI models learn from existing data, they may learn, and possibly amplify, problematic biases, based for example, on race, gender or religion. At the same time Google (undated a) emphasised that there were some "AI applications we will not pursue" (para. 1), and listed four such applications areas. These areas covered technologies that cause or are likely to cause overall harm; weapons or technologies whose main purpose was to injure people; technologies that gather or use information for surveillance in violation of internationally accepted norms; and technologies whose purpose contravened acceptable principles of international law and human rights.

Microsoft argued that explainability was important in that it was seen to help scientists, auditors and decision makers ensure that decisions about AI systems can be reasonably justified, and that such systems comply with company policies, industry standards and government regulations. It was suggested, for example, that a data scientist should be able to explain to a

stakeholder how a certain level of accuracy was achieved and what influenced this outcome. In focusing on trust, Accenture (2023, para. 27) argued that "every company using AI is subject to scrutiny", that ethical concerns "where companies amplify their responsible use of AI through public relations while partaking in unpublicized gray-area activities, is a regular issue" and that "unconscious bias is yet another", but that "responsible AI is an emerging capability aiming to build trust between organizations and both their employees and customers".

Microsoft emphasised that accountability is an essential pillar of responsible AI. Here the company's argument was that the people who design and deploy AI must be accountable for their actions and decisions, not least as more autonomous systems are being developed. Accenture (2023, para. 28) argued that "data privacy and the unauthorised use of AI could be detrimental both reputationally and systematically", and "that companies must design confidentiality, transparency and security into their AI programs at the outset and make sure data is collected, used, managed and stored safely and responsibly".

In addressing control, Accenture (2023, para. 30) argued that "machines don't have minds of their own, but they do make mistakes", and that "organizations should have risk frameworks and contingency plans in place in the event of a problem", and that it is important to "be clear about who is accountable for the decisions made by AI systems, and define the management approach to help escalate solutions when necessary". In outlining recommended practices on a human-centred design approach, Google argued that the way users experienced AI systems was essential in assessing the impact of their predictions and decisions. Here, the focus was seen to be on "designing features with appropriate disclosures built-in", where "clarity and control is central to the user experience" (Google, undated b, para. 4).

The selected companies provided specific examples to demonstrate their AI responsibilities. Under the banner "Responsible AI with Google Cloud", for example, Google (undated b, para.10) argued that rigorous evaluation was "a critical component of building successful AI". Here, the company reported that two diverse review bodies undertook "deep ethical analysis and risk opportunity assessments" to drive Google Cloud's alignment with the company's AI principles. Under the banner "Responsible AI with TensorFlow", Google (undated b, para.11) showcased a consolidated toolkit for third party developers on TensorFlow, designed to help to build machine learning fairness, interpretability, privacy and security into their models. Microsoft (2023b, paras. 2 and 3) offered details of its "Responsible AI Standard" - the company's internal playbook for responsible AI, which "shapes the way in which we create AI systems, by guiding how we design, build and test them", and of its "Responsible AI Impact Assessment Template".

Discussion

This paper has provided some insights into the ways in which four major ICT companies have identified, and claimed to be publicly addressing, their responsibilities towards AI, and while the paper did not allow any comparable analysis, a number of common threads can be identified. The four companies recognised that the deployment of AI raised a number of issues and challenges, and in looking to address such challenges, the four companies emphasised their commitment to a number of principles that they claimed guided their approach to their AI responsibilities. That said, the paper revealed that all four companies prefaced their approach to their AI responsibilities with positive outline appraisals of the benefits of AI, both to companies and customers. In taking a positive approach to AI and focussing on its benefits at a corporate and individual level, major companies might be seen to be effectively playing down its potentially negative impacts. As such, the approach taken by the selected ICT companies towards CDR and AI could be seen at best as part of a major corporate marketing/public relations exercise, and at worst, as ethics washing, namely feigning ethical consideration, designed to improve how companies are perceived by stakeholders.

More generally, the main focus of the four company's approaches to their AI responsibilities is largely centred on social and technology issues. In outlining their social responsibilities – fairness, for example - inclusion and avoiding bias loom large, but little, or no, attention, is paid to environmental issues, and more particularly, to climate change. Climate change has been described by the United Nations (2023, para. 2) as "the defining issue of our time", and it may have fundamental social impacts, including the wholesale destruction of homes and communities, the loss of livelihoods, population migration and forced displacement, and the loss of cultural identity. Here, the poorest sections of society and most vulnerable social groups may bear the brunt of the changes. While such essentially social issues do not feature in the four selected companies' approaches to their AI responsibilities, AI can be seen to offer both a major opportunity to mitigate climate change, and to be a cause of such change.

On the one hand, the United Nations Environment Programme (2022, para. 1) outlined "how artificial intelligence is helping to tackle environmental challenges", namely "a triple planetary crisis of climate change, nature and biodiversity loss, pollution and waste". Further, the United Nations Environment Programme (2022, para. 2) claimed that though "more climate data is available than ever before……. how that data is accessed, interpreted and acted on is crucial to managing these crises", and that "one technology that is central to this is AI". More specifically, AI is seen to have a vital role to play, for example, in helping to measure greenhouse gas emissions, to reduce the impact of such emissions, to remove existing emissions from the atmosphere, and to improve hazard forecasting for both long term events, such as rises in sea levels, and for short term extreme events, such as hurricanes.

On the other hand, the United Nations Environment Programme (2022, paras. 16 and 17) warned that while "data and AI are necessary for enhanced environmental monitoring, there is an environmental cost to processing this data", not least that "the ICT sector generates about 3-4 percent of emissions, and data centres use large volumes of water for cooling". From an arguably more critical academic perspective, Nost and Colven (2022, p. 23) mapped out two case studies to "show that environmental and climate crises are grist for tech solutions and find that many climate AI actors are interested in it for surveillance, greenwashing, and commodifying algorithms". While Wynsberghe (2021, p. 217) suggested that the environmental costs of AI could not be ignored, looking more positively towards the future, he emphasised the need to "develop AI that is compatible with sustaining environmental resources for current and future generations".

There are also potential problems for some stakeholders in fully understanding how companies discharge their technological responsibilities to AI. In a study of the implications of AI for the

concept of responsibility in the human rights arena, the Council of Europe (2019, p. 9), highlighted "the power asymmetry between those who develop and employ AI technologies, and those who interact with and are subject to them". Here the argument was that while digital service providers, and linked companies that utilise AI systems, can acquire very detailed data about the users of their services, which they can mine to generate predictions about user traits, tastes and preferences with considerable accuracy, the users themselves typically do not understand the complexities of the digital technologies that they use.

Further, the Council of Europe (2019, p.9) suggested that this asymmetry not only expands the threat of potential exploitation, but that such risks are exacerbated by the capacity of AI technologies to operate at unprecedented speed and scale, which can generate new challenges for society. In a related argument, Scorici *et al.* (2022, para. 1**)** identified a gap in the relationship "between humans and AI, triggered by commercial interests, making use of AI through advertisement, marketing, and corporate communications". More worryingly, Scorici *et al*. (2022), went on to outline the deceptive use of AI enabled machines, aimed at - intentionally or unintentionally - misleading stakeholders and the general public, about the true capabilities of AI. The authors concluded by calling for companies to address the power asymmetries associated with deployment of AI, and to take responsibility where stakeholders have been, or might be, misled.

Conclusion

This exploratory paper has suggested that while the digital technologies bring a wide range of new business benefits and opportunities, as an ever-growing number of companies adopt these technologies, so they may have to face, and increasingly address, several new sets of responsibilities. AI is arguably seen as the most powerful of the digital technologies, and it can pose particularly testing challenges for CDR. This paper offers some insights into how four major ICT companies are publicly addressing their AI responsibilities. These responsibilities are increasingly being captured in the concept of CDR, but how companies are addressing CDR has received only limited attention in the academic literature. As such, the paper can be seen to contribute to helping to fill this gap in the academic literature.

At the same time, the themes identified as characterising the selected companies' public approach to their AI responsibilities can also be seen in the light of some of the literature reviewed earlier in the paper. While Wirtz *et al.* (2022a) claimed that CDR had not been explored in a service context, the paper provides some illustrations of how four large companies - that are technology providers but also operate in the service sector - are addressing CDR, in the context of AI. More specifically, the paper offers illustrations of how these companies are looking to acknowledge their social and technological responsibilities associated with AI, whilst also pointing out that both environmental responsibilities and the asymmetry of power between developers and users receive scant attention from the selected companies. While all four companies claim to work within the law in all jurisdictions in which they operate, no detailed attention is paid to the role of the state in governance procedures. The paper can also be seen to contribute to the embryonic attempts to conceptualise CDR, in that some of the themes presented in the AI responsibility narrative map onto the proposed norms identified by Mihale-Wilson *et al.* (2022), and in that some of the underlying rationales for voluntary disclosure theory, signalling theory and legitimacy theory, suggested by Bonson *et al.* (2023), also resonate.

That said, the paper does have a number of limitations, not least in that it is largely based on material from just four companies, and in that this material is drawn exclusively from Internet sources. At the same time, while the AI responsibility themes identified in the paper offer illustrations of how the four selected companies are publicly addressing CDR, they do not provide a comprehensive picture, or a detailed analysis, of the development and workings of these themes in practice. Nevertheless, the paper can be seen to provide a platform for future research agendas, which may help to lay the foundations for a more comprehensive review of CDR. As Yahaya *et al*. (2023, p.6) conclude "the AI revolution is upon us and the way we live, do businesses, run economies and countries are already being transformed. Additionally, the continuous investment and research focus on further development of artificial intelligence shows that the future of individual lives, businesses and economies will continuously be influenced by numerous everyday artificial intelligence functions". Future research in this area of study will thus inevitably grow, and could include, for example, empirical work to explore how a range of companies across the business spectrum are addressing CDR, and if, and how, a variety of stakeholders have been involved in that process. A detailed examination of the role of the state in corporate AI governance procedures, and the locus of power between companies and the state, also merits research attention. In the light of the findings of such empirical research endeavours, it will also be important to test, and refine, existing theoretical frameworks to conceptualise CDR within business and social contexts.

References

1. Accenture (2023). *Artificial Intelligence*. Available online: https://www.accenture.com/gb-en/insights/artificial-intelligence-summary-index#:~:text=At%20Accenture%2C%20we%20define%20%E2%80%9Cresponsible,and%20scale%20AI%20with%20confidence.&text=Every%20company%20using%20AI%20is%20subject%20to%20scrutiny (accessed on 3 March 2023).
2. Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. doi: http://dx.doi.org/10.1080/1364557032000119616
3. Armstrong, R., Hall, B. J., Doyle, J., & Waters, E. (2011). "Scoping the scope" of a Cochrane review. *Journal of Public Health, 33*(1), 147–150. doi: http://dx.doi.org/10.1093/pubmed/fdr015
4. Bonson, E., Bednarova, M., & Pera, D. (2023). Disclosures about algorithmic decision making in the corporate reports of Western European companies. *International Journal of Accounting Information Systems, 48*. Available online: https://www.sciencedirect.com/science/article/pii/S1467089522000483 (accessed on 21 February 2023).
5. Borenstein, J. & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. *AI and Ethics, 1*, 61-65.
6. Constantinescu, M., Voinea, C., Uszkai, R., & Vica, C. (2021). Understanding responsibility in Responsible AI. Dianoetic virtues and the hard problem of context. *Ethics and Information Technology, 23*, 801-814.
7. Corporate Digital Responsibility (2022). *What is CDR?* Available online: https://corporatedigitalresponsibility.net/cdr-manifesto-english (accessed on 21 February 2023).

8. Council of Europe (2019). *Responsibility and AI*. Available online: https://rm.coe.int/responsability-and-ai-en/168097d9c5 (accessed on 5 March 2023).

9. Danaher, W. (2022). Tragic Choices and the Virtue of Techno-Responsibility Gaps. *Philosophy and Technology, 35*. Available online: https://link.springer.com/article/10.1007/s13347-022-00519-1 (accessed on 10 April 2023).

10. Djeffal, C., Siewert, M.B., & Wurster, S. (2022). Role of the state and responsibility in governing artificial intelligence: A comparative analysis of AI strategies. *Journal of European Public Policy, 20*(11), 1799-1821.

11. Elliott, K., Price, B., Shaw, P., Spiliotopoulos, T., Ng, M., Coopamootoo, K., & van Moorsel, A. (2021). Corporate Digital Responsibility. *Society, 51*, 179-188.

12. Enholm. I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial Intelligence and business Value: a Literature Review. *Information Systems Frontiers, 24*, 1709-1734.

13. Finfgeld-Connett, D. (2014). Use of Content Analysis to Conduct Knowledge-Building and Theory-Generating Qualitative Systematic Reviews. *Qualitative Research, 14*, 341–352.

14. Flyvbjerg, B. (2006). Five Misunderstandings about Case-Study Research. *Qualitative Inquiry, 12,* 219–245. doi:10.1177/1077800405284363

15. Google (undated a). *Artificial Intelligence at Google: Our Principles*. Available online: https://ai.google/principles/ (accessed on 12 April 2023).

16. Google (undated b) *Responsible AI Practices*. Available online: https://ai.google/responsibilities/responsible-ai-practices/ (accessed on 22 February 2023).

17. Haenlein, M. & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review, 61*(4), 5-14. Available online: https://doi.org/10.1177/0008125619864925

18. Hanneke, R., Asada., Y., Lieberman, L., Neubauer, L. & Fagan, M. (2017). The Scoping Review Method: Mapping the Literature in Structural Change Public Health Interventions. *Department of Public Health Scholarship and Creative Works, 94*, 1-14. SAGE Research Methods Cases Part 2. Available online: https://digitalcommons.montclair.edu/public-health-facpubs/94 (accessed on 12 April 2023). Doi: http://dx.doi.org/10.4135/9781473999008

19. Herden, C.J., Alliu, E., Cakici, A., Cormier, T., Deguelle, C., Gambhir, S., Griffiths, C., Gupta, S., Kamani, S. R. Kiradi, Y-S., Kispataki, M., Lange, G., de Mato, L.M., Moreno, Nunez, H.A., Pilla, V., Raj, B., Roe, J., Skoda, M., Song, Y., Ummadi, P.K., & Edinger-Schons, L.M. (2021). Corporate Digital Responsibility. *Sustainability Management Forum, 29*, 13-29.

20. IBM (2022). *AI ethics in action: An enterprise guide to progressing trustworthy AI.* Available online: https://www.ibm.com/downloads/cas/4DPJK92W (accessed on 24 February 2023).

21. IBM (undated). *AI Ethics.* Available online: https://www.ibm.com/topics/ai-ethics#:~:text=IBMs%20AI%20Ethics&text=IBM%20seeks%20to%20ensure%20that,compliance%20with%20ethical%20AI%20standards (accessed on 24 February 2023).

22. International Telecommunications Union (2020). *Addressing the dark side of AI.* Available online: https://aiforgood.itu.int/event/addressing-the-dark-sides-of-ai/ (accessed on 26 February 2023).

23. Jyoti, R. (2022). *Scaling AI/ML Initiatives: The Critical role of Data* Available online: https://www.snowflake.com/wp-content/uploads/2022/03/Scaling-AI-ML-The-Critical-Role-of-Data-.pdf (accessed on 20 February 2023).

24. Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021). Corporate Digital Responsibility. *Journal of Business Research, 122*, 875-888.

25. Microsoft (2023a). *What is Microsoft's Approach to AI?* Available online: https://news.microsoft.com/source/features/ai/microsoft-approach-to-ai/ (accessed on 12 April 2023).

26. Microsoft (2023b). *Responsible AI*. Available online: https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1:primaryr6 (accessed on 23 February 2023).

27. Mihale-Wilson, C., Hinz, O., van der Aalst, W., & Weinhardt, C. (2022). Corporate Digital Responsibility. *Business and Systems Information Engineering, 64*, 127-132.

28. Mikalef, P., Conroy, K., Lundstrom, J. E., & Popovic. A. (2022). Thinking responsibly about responsible AI and the dark side of AI. *European Journal of Information Systems, 32*(3), 257-269.

29. NetApp (2023). *What is Artificial Intelligence?* Available online: https://www.netapp.com/artificial-intelligence/what-is-artificial-intelligence/#:~:text=Today%2C%20the%20amount%20of%20data,of%20all%20complex%20decision%20making (accessed on 20 February 2023).

30. Nost, E. & Colven, E. (2022). Earth for AI: A Political Ecology of Data-Driven Climate Initiatives. *Geoforum, 130*, 23-34. Available online: https://www.researchgate.net/publication/358662265_Earth_for_AI_A_Political_Ecology_of_Data-Driven_Climate_Initiatives (accessed on 4 March 2023).

31. Piquard, A. (2023). Elon Musk and hundreds of experts call for 'pause' in AI development. *Le Monde*. March 29 2023. Available online: https://www.lemonde.fr/en/economy/article/2023/03/29/elon-musk-and-hundreds-of-experts-call-for-pause-in-ai-development_6021147_19.html (accessed on 11 April 2023).

32. Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme. Lancaster University. Available online: https://www.researchgate.net/publication/233866356_Guidance_on_the_conduct_of_narrative_synthesis_in_systematic_reviews_A_product_from_the_ESRC_Methods_Programme (accessed 12 March 2023).

33. PWC (2023). *Corporate Digital Responsibility and Digital Ethics*. Available online: https://www.pwc.de/en/sustainability/corporate-digital-responsibility-and-digital-ethics.html (accessed on 23 February 2023).

34. Scorici, G., Schultz, M.D., & Seele, P. (2022). Anthropomorphization and beyond: conceptualising humanwashing of AI-enabled machines. *AI and Society.* Available online: https://doi.org/10.1007/s00146-022-01492-1 (accessed on 6 March 2023).

35. Tigard, D. (2021). There Is No Techno-Responsibility Gap. *Philosophy and Technology, 34*, 589-607.

36. United Nations (2022). *Climate Change*. Available online: https://www.un.org/en/global-issues/climate-change (accessed on 5 March 2023).

37. United Nations Environment Programme (2022). *How artificial intelligence is helping to tackle environmental challenges.* Available online: https://www.unep.org/news-and-stories/story/how-artificial-intelligence-helping-tackle-environmental-challenges (accessed on 5 March 2023).

38. Wirtz, J., Hartley, N., Kunz, W. H., Tarbit, J., & Ford, J. (2021). Corporate Digital Responsibility at the Dawn of the Digital Revolution. *SRN Electronic Journal.* Available online: https://www.semanticscholar.org/paper/Corporate-Digital-Responsibility-at-the-Dawn-of-the-Wirtz-Hartley/7d0809b389d45b5a4dc045b31e53de91ec014167 (accessed on 12 March 2023).

39. Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2022a). Corporate Digital Responsibility in Service Firms and Their Ecosystems. *Journal of Service Research*. Available online: https://web2-bschool.nus.edu.sg/wp-content/uploads/media_rp/publications/t0rkf1666776069.pdf (accessed on 20 February 2023).

40. Wirtz, J., Tarbit, J., Hartley, M. and Kunz, W. (2022b). Corporate digital responsibility: Dealing with ethical, privacy and fairness challenges of AI. *Journal of AI, Robotics and Workplace Automation, 1*(4), 1-4.

41. Yahaya, M., Umagba, A., Obeta, S., & Maruyama, T. (2023). Critical Evaluation of the Future Role of Artificial Intelligence in Business and Society. *Journal of Artificial Intelligence, Machine Learning and Data Science*, *1*(1), 1-9. Available online: https://urfpublishers.com/open-access/critical-evaluation-of-the-future-role-of-artificial-intelligence-in-business-and-society.pdf (accessed on 12 March 2023).

42. Zoshak, J. & Dew, K. (2021). Beyond Kant and Bentham: How ethical Theories are being used in Artificial Moral Agents. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Available online: https://doi.org/10.1145/3411764.3445102 (accessed on 6 April 2023).