Revising the 'science of the organisation':
Theorizing AI agency and actorhood

Mark van Rijmenam
University of Technology Sydney, Australia
Mark.vanRijmenam@alumni.uts.edu.au (corresponding author)

Danielle Logue
University of Technology Sydney, Australia

Forthcoming in
Innovation: Organization & Management
August 2020

DOI: 10.1080/14479338.2020.1816833

Abstract

Artificial intelligence is a central technology underpinning the fourth industrial revolution, driving dramatic changes in contemporary cyber-physical systems and challenging existing ways of theorizing organizations and management. AI agency and the rise of the artificially intelligent agent is both fundamentally different and yet increasingly similar to human agency in terms of intentionality and reflexivity. As 'Child AI' emerges—AI that is created by other AI—the early human design and interaction becomes increasingly distant and removed. These developments, while seemingly futuristic, changes the human-technology interface through which we organise. In this essay we explore understandings of AI agency, capability and governance, and present implications for organizational theorising in sociomateriality, actor-network theory, institutional theory and the behavioral theory of the firm. We contribute to a growing and reflexive research agenda that can accommodate and regenerate theorizing around this significant technological advancement.

Introduction

The fourth industrial revolution is an era that is and will be characterised by cyber-physical systems, underpinned by developments in artificial intelligence (AI) and a corresponding rising interest in and concern over events at the human-technology interface. While examples now abound of computers' strategic game-playing ability due to their processing and algorithmic ability to analyse vast volumes of data (Agrawal et al., 2019; von Krogh, 2018), more recent developments represent a step change in the self-learning capacity and agency of AI. In 2016, Google's Deepmind developed AlphaGo, an AI actor that learned an abstract strategy board game (Go) with a far more expansive range of moves than chess (Wang et al., 2016). A year later, Deepmind developed AlphaGo Zero, simply by playing against itself (Silver & Hassabis, 2017). Within three days, it surpassed the original AlphaGo, an algorithm that had beaten 18-time world champion Lee Sedol. By 2020, AI has again further advanced.

Today, companies such as Deepmind and Open AI (the AI research organisation originally founded by Elon Musk) demonstrate the capabilities of AI actors in complex games based on deep neural networks, trained directly from raw game data via both supervised and reinforcement learning, that also require collaboration between AI players, and result in scores in computer games 'above the human baseline', i.e. outperforming any human players (Deepmind, 2020). In addition, researchers from Google, building on earlier research where they managed to have AI create AI (so-called 'Child AI') (Le & Zoph, 2017), have now developed AI that can improve generation after generation, without any human involvement. Replicating decades of AI research in a matter of days (Real et al., 2020), the program, called AutoML Zero, can build AI agents without human input, using only basic mathematical concepts, eventually leading to entirely new AI capabilities.

These examples highlight the rapid evolution of artificially intelligent agents, with some even arguing that AI will eventually consist of entirely novel and unrecognisable forms of intelligence (Armstrong et al., 2012; Bostrom, 2014). With AI advancing, the ability of AI agents to act independently of human oversight, respond to their environment and interact with other machines increases (potentially without a human to check and validate decisions and intentionality). We now have a situation whereby artificially intelligent agency (AI agency) goes on to produce more artificially intelligent agency and so on, becoming further and further removed from the initial human design or interaction.

Historically, existing theories of technology's role in organizing processes have been largely adequate to explain consequent impact and changes in work, management and organisation (Leonardi & Barley, 2010; Orlikowski & Scott, 2008; Zammuto et al., 2007). Yet, with the introduction and rapid advancement of artificial intelligence, these technological processes are becoming increasingly 'intelligent' and autonomous, appearing intentional and indistinguishable from human efforts, and increasingly outperforming such human effort in a variety of tasks and cognitive acts. This ability to learn and act autonomously distinguishes intelligent technological actors from most technologies historically used in organisations, resulting in new forms of organizing and challenging existing conceptualisations of technology in organisational theory (Bailey et al., 2019; von Krogh, 2018; Faraj et al., 2018; Baum & Haveman, 2020; Dove & Fayard, 2020). Thus, in this essay, we argue that this evolution of AI agency and actorhood is more than just a technological development of which to both be aware and make sense—it also presents serious challenges for organisational theorizing and potential revisions to our assumptions, leading to a new science of the organisation.

In the following sections, we first present both our definitions and starting points for understanding AI and artificially intelligent agency, and propose that organisational scholars, while increasingly attentive to artificial intelligence, robots and algorithms generally (Benaich & Hogarth, 2019; Fleming, 2019; Flyverbom, 2019; Baum & Haveman, 2020; Kellogg et al., 2020), need to further consider the theoretical challenges and implications of AI agency and its near-future manifestations. To support this argument, we begin by bringing to bear sociomateriality as a theoretical lens on AI agency, followed by considerations of actor-network theory, institutional theory and the behavioural theory of the firm to further demonstrate how this step-change in AI agency challenges many of our theoretical lenses. We conclude with the beginnings of a research agenda on how these recent advancements in applications and capabilities of AI will affect organisation design, strategic management, decision-making, power and other issues of interest to management and organisational theorists.

Artificially Intelligent agency

The financial impact of AI on the global economy is estimated to reach US$15.7 trillion by 2030 (Rao & Verweij, 2017), with 40% of jobs expected to be lost due to AI (Lee, 2018), and with global venture capital investment in AI more than US$27 billion in 2018 (Benaich & Hogarth, 2019). Already, 37% of organisations are experimenting with AI in some form (Gartner, 2019), and in recent years the number of AI-related patents has also surged (Tseng & Ting, 2013). Between 2013-2016, the number of patent filings related to deep learning, increased from 118 patents in 2013 to 2399 in 2016, the equivalent of a 175 percent average annual growth rate (AAGR), while patents for all technologies only grew with 33 percent, or a 10 percent (AAGR) over the same period (World Intellectual Property Organisation, 2019).

In defining AI and AI agents, numerous definitions of AI emphasise its function as a computer-assisted system for task inputs, processes and outputs (Carbonell, 1970; Kandel, 1991; Norman & Draper, 1986). Other definitions focus on its role in completing tasks that usually require human intelligence, in areas such as visual perception and speech recognition (Radford, 2019; Russell & Norvig, 1995; von Krogh, 2018). We simply draw on a commonly used definition of AI as a 'system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation' (ESCP Europe Business School, Paris, France, as cited in Haenlein & Kaplan, 2019, p. 1). The recent examples of game-playing AI mentioned earlier show how AI agents cannot only simulate intelligent behaviour, as in playing a computer game, but can also outperform human intelligence, albeit still in a narrow domain. This simulation and outperformance of intelligent behaviour can be done using various computing techniques, which include, but are not limited to, machine learning, deep learning, reinforcement learning, natural language processing and image recognition. This brings us to our definition of AI agents: *artificially intelligent actors that have the ability to imitate, and outperform, human intelligence, act upon their own, distinct from and without further human intervention*. Child AI[1] is defined as *an AI agent developed by an AI agent, without any human intervention*. Consequently, AI agents can address a wide array of problems, including perception, reasoning, knowledge, planning and communication, and across research paradigms such as the symbolic (applying logic- and knowledge-based tools in areas such as robotic automation and expert systems); statistical (focusing on probabilistic methods and machine learning, in areas such as decision networks, natural language processing and neural networks); and sub-symbolic (including intelligent search and optimisation, and embodied intelligence approaches, working in areas such as evolutionary algorithms and autonomous systems) (Corea, 2019). We note here the distinction between AI and algorithms, defining an algorithm as a set of rules or instructions (a pre-set, rigid and coded recipe), i.e. *If This Then That Statements* albeit a lot more complex, to be followed by a computer to solve a certain problem when it encounters a trigger. AI, on the other hand, is a group of algorithms working together to autonomously solve a certain problem, without having to 'wait' for a certain trigger. Based on learned inputs and data, AI can change or create new algorithms and AI agents.

[1] From here on, we will refer to AI when talking about Child AI, as in the end, also Child AI is AI.

What we observe today is AI agents having the capacity to change their behaviour (Curchod et al., 2019), and collaborate, strategize and make decisions independently and autonomously, thereby altering context without being subject to further human action. Significantly, AI agents can (by definition) act on its own: it showcases intelligent behaviour, it has goals and it can reason and monitor its own behaviour (Bostrom, 2014). It can increasingly reproduce and evolve without human action (Abadi & Andersen, 2016; Lewis et al.,, 2017; Le & Zoph, 2017; Missinato, 2018; Real et al., 2020). Such capabilities and capacity, we argue, raise many implications for organisational theorizing.

AI Agents and Challenges for Organisational Theorizing

In his classic work, *The Sciences of the Artificial* (1996), Herbert Simon identified digital elements as artificial things that are synthesised by humans and, therefore, not natural. He argued that to better understand how the artificial will behave, a process of simulation may be used, because the behaviour of computers is governed by simple laws (algorithms) where the complexity results from the environment (data). However, in the past years, highly complex neural networks have transformed those simple laws into complex AI agents. Such technological developments raise questions—technical, philosophical, and organisational—as to the ultimate capability, control, governance and morality of such developments, especially when AI becomes more intelligent over time. Organisationally, we identify three areas of concern: objectivity, opacity and ordering. Objectivity, with calls for more human oversight as we observe automated forms of analysis and decision often amplifying inequality and bias; opacity and transparency whereby artificial agency enables both governance and mass surveillance; and ordering, whereby AI agents become involved in the ordering of social life and the institutional conditions that are at stake when AI is introduced. For example, Bell (2017, 2018) asks key questions about autonomy, assurance and agency: How do we design for an autonomous world? How do we preserve our safety and values? Further, as our familiar methods of interfacing with computing systems (e.g. screens and keyboards) are superseded, what consequences will arise when AI-enabled systems surround us, sensing and responding to us, and making decisions (Shrestha et al., 2019)? To this end, Himma (2009) asked if it is indeed possible to produce artificially intelligent moral agents, while Bostrom (2014) discussed ways of responding to AI that becomes 'super intelligent', being fundamentally different than human intelligence[2], surpassing it and transforming societies in unimaginable ways (Van Rijmenam, 2019).

We continue in a broad tradition of researching the impact of new technologies on organisational theorizing and the relationship between organisational design, structure, performance and technology (Bostrom & Heinen, 1977; Emery & Trist, 1960; Woodward, 1965). Early waves of such work focused predominantly on manufacturing technology, with later research sought to include a variety of other technologies (Glisson, 1978; Perrow, 1967). For example, Huber (1990), argued for a revision of existing organisation design theories because the first wave of advanced information technologies changed the nature of organisational design, intelligence and decision-making. These information technologies (e.g. digital columns, records, numbers and algorithms) were perceived by some as material (Dourish & Mazmanian, 2011; Mazmanian et al., 2014), whereas others argued that, since these digital elements have no weight and lack any spatial mode of being, they should be considered

[2] AI will be fundamentally different from human intelligence. Intelligence is 'the complex expression of a complex set of principles' (Yudkowsky 2007, 389), which consists of multiple interdependent subsystems linked to each other. Intelligence exists due to evolution and enables humans to model, predict and manipulate reality. It enables us to reason backwards and forwards from a mental image, and reason regarding desired future outcomes (Yudkowsky, 2007). Evolution created intelligence, but evolution does not possess this foresight. In fact, this evolutionary process is an unintelligent process and has resulted in flaws in human intelligence (Yudkowsky, 2008). Due to various constraints (such as food availability and trade-offs with other organs or biological materials) our brains may not have evolved in the most optimised way (Armstrong et al., 2012). Since AI is developed by (artificially) intelligent actors with foresight capabilities that evolution does not possess (Yudkowsky, 2007), and uses materials and processes better suited for intelligence, it is likely that AI consist of new forms of intelligence unfamiliar to humankind today (Armstrong et al., 2012; Bostrom, 2014).

immaterial (Faulkner & Runde, 2012; Kallinikos, 2012), and that digital elements are artificial, synthesised and unnatural things (Simon, 1996).

In this essay, we make a step-change from existing theorizing on AI and consider AI in its likely future manifestation, perhaps 10, 20 or 30 years from now, when AI is truly intelligent, and evolved to a state of so-called Artificial General Intelligence (AGI). AGI refers to AI systems having autonomous self-control and self-understanding and the ability to learn new things to solve a wide variety of problems in different contexts (Goertzel & Pennachin 2007; Bostrom, 2014). If one takes a more forward-looking perspective and envisions AI as what it undoubtedly will become in the future, an artificially intelligent agent that has the power to outperform human agents on most if not all domains, it becomes clear that existing theories will be challenged as we attempt to understand this changing human-machine interface. In what follows we describe the implications for several theoretical lenses given their dominance in theorizing innovation, organisation and management: sociomateriality, actor-network theory (ANT), institutional theory and the behavioural theory of the firm, to begin informing a broader research agenda on AI agency and organisational theorizing and to also show the far-reaching implications theorizing AI will have across different perspectives.

AI agency and sociomateriality

Historically, the notion of agency has been viewed as a strictly human capacity (Sørensen & Ziemke, 2007). However, as AI becomes more advanced, discussions are emerging as to how to conceptualise AI agency and its intentionality (Johnson & Noorman, 2014; Johnson & Verdicchio, 2019; Murray et al., 2020). Our starting point is that technology is not only embedded, shaped and informed by socio-organisational forces but also influences those forces (Fleming, 2019; Orlikowski, 2000; Orlikowski & Robey, 1991; Weißenfels et al., 2016), drawing us to sociomateriality as an initial and rich theoretical lens through which to interrogate AI agency as it describes the 'space in which the social and the material become entangled' (Orlikowski, 2009). Leonardi et al. (2012, p. 42) define social agency as 'coordinated human intentionality formed in partial response to perceptions of a technology's material agency' and material agency as 'ways in which a technology's materiality acts'. Emirbayer and Mische (1998), described human agency as 'the temporally constructed engagement by actors of different structural environments. which, through the interplay of habit, imagination, and judgment, both reproduces and transforms those structures.' Material agency is activated as humans approach technology with particular intentions and decide which elements of its materiality to use at a given time' (Leonardi *et al*. 2012, p. 42). Yet we find these definitions insufficient for understanding AI agency. Artificially intelligent entities can exercise agency through their performativity, that is, by doing things that are outside the control of users or other artificial or human agents and when agents' actions materialise through their intentionality and reflexivity, objectives can be achieved (Muñoz & Encinar, 2014). The more AI gains autonomy and agency, the more it will be responsible for its own development (Armstrong, 2017; Turchin & Denkenberger, 2020). In the (near) future, this can result in AI showcasing computational elements that can 'make them agents the way humans are' (Bostrom, 2014; Omohundro, 2016; Johnson & Verdicchio, 2019, p. 645). Given this, and within the tradition of sociomateriality, we would define AI agency as *coordinated artificially intelligent intentionality formed in partial response to perceptions of human agency, material agency and/or other AI agency*.

Orlikowski's (2009) earlier work assists in the theorisation of AI agency, by describing how technology is the result of continuous interactions between human actors, actions, choices, social histories and institutional contexts; therefore, its materiality is socially defined and produced and only relevant to people engaging with it (Orlikowski, 2009). She highlighted that technology will produce certain identifiable impacts on organisations (Orlikowski, 2009), which definitely applies to AI and its influence on firms (Murray et al., 2020). Orlikowski and Robey (1991) argued that information technology has social and material properties; it is constructed by human actions and objectified by institutionalisation. In this way, information technology offers a framework for human actors to understand their world, offering a medium for the construction of their social reality and contributing to human actions via objectifying knowledge and assumptions (Orlikowski & Robey, 1991). However, the integration of artificial intelligence (including deep learning and complex neural networks) into organisations and, specifically, the scale and production capacity of such AI agents—that is, the independent and autonomous ability to produce additional AI and further AI agency—raises significant implications. For example, when no 'social' is involved, as in the case of AI creating AI (Le & Zoph, 2017), does the entanglement of social and material still take place? How can we conceptualise this interaction when no human actions are involved in the creation of technology but, rather, a (Child) AI agent that has been created by technology itself[3]? For example, one traditional understanding holds that the use of technology is influenced by the understanding of the user (Orlikowski, 2000). In their use of technology, humans are influenced by its materiality, inscribed by both the designer and previous users. Therefore, previous use affects current use, even in ways not originally intended by designers; human actors might continue to use it in that way for any number of reasons (e.g. due to corporate pressure, unavailability of staff, user expectations or increased knowledge about that particular technology). However, in the case of AI created by AI, this is no longer the case. AI agents are not bound by issues such as corporate pressure or unavailability of staff (i.e. AI does not get sick, is always available and sticks to the reason why it was developed and, does not give in to corporate pressure to

[3] It is true that almost all AI created to date is biased. After all, AI is often trained with biased data and developed by biased humans (O'Neil, 2016). This means that the social is very much involved in the creation of such AI. However, we are now seeing developments of AI being trained without (biased) data at all (Deepmind, 2018, 2020) or created without (biased) developers (Le & Zoph, 2017). Upcoming developments, such as self-supervised learning could become a technique that would create data-efficient AI systems (LeCun, 2020)

change its behaviour). AI also eliminates wrong, unsuccessful behaviour if it does not contribute to achieving its ultimate goal (Bostrom, 2014), i.e. AI always aims to improve its outcome[4]. Existing structurational models (that is, models that consider technology a product of human action, both in its original physical construction, and in its later social construction through use) offer little help in this regard because they assume that technology stabilises over time (meaning that human actions do not refine and modify technology) (Orlikowski & Robey, 1991; Orlikowski, 2000); in a world of transfer learning, self-learning algorithms using reinforced feedback loops, and AI agents creating AI this is no longer the case.

AI agency also challenges Orlikowski's (2000) valuable concept of 'technology-in-practice': that technology can be constructed with certain materials and assumptions, but that it only structures human action when it is used by said action. When an actor decides to use a technology, they also decide how to interact with it; these interactions can change in different times and places (i.e. driving a car in a different country with different rules). However, physical properties result in boundary conditions that govern how to use an artifact; the more said artifact is integrated into a system, the narrower the alternative uses it has. And so how can we understand the 'matter' of AI agency? Fine-grained material characteristics (bits and bytes) makes it very pervasive and ambiguous at a granular level previously unthinkable, such that technology includes all levels of material from completely artificial environments to miniaturised devices or even quantum mechanics (Leonardi et al., 2012). As such, the materiality of technological artifacts has both material features, which might not be directly observable as is the case with software, and material consequences. On the other hand, digital elements have no weight and lack any spatial mode of being, so they should be considered

[4] If an AI has been trained with biased data, AI will still always improve itself from an AI perspective. This means, it is becoming better at the objective it was given (for example, hiring the right candidate) but it might no longer be seen as an improvement from a human perspective (it only hires male candidates as the biased training data showed that males where hired more often in the past). As O'Neil (2016) clearly showed, an AI that discriminates has not been built correctly by the developer, but it will still always improve itself over time (Bostrom, 2014).

immaterial (Dourish & Mazmanian, 2011; Faulkner & Runde, 2012; Kallinikos, 2012; Mazmanian et al., 2014; Orlikowski, 2007). Consequently, while AI's form and function is increasingly becoming detached from matter (Leonardi et al., 2012), it does so in a Heideggerian sense (Riemer & Johnston, 2017), withdrawing from materiality and becoming increasingly 'invisible' to humans[5] i.e. AI consist of only bits and bytes and thus itself can be considered immaterial (Faulkner & Runde, 2019).

AI agency also seems to challenge the concept of entanglement. According to Leonardi (2013), technological artifacts are created by social action, which in turn shape human action. Within this entanglement, the material influences the social and vice versa and all organisational aspects are bounded by the material (Orlikowski 2007). Within this agential realism approach, the social and the material are interpenetrated and entangled, while a critical realist approach sees the social and the material as separate, becoming entangled by means of action (Tunçalp 2016). Indeed, when applied to artificially intelligent agents, they have the capacity to act autonomously in response to human and material agency. AI is social because it is developed by humans; however, it is also nonsocial because, increasingly, AI artifacts can now be created by other AI artifacts, without any human involvement (Le & Zoph, 2017; Real et al., 2020). As Ullman describes:

> *when programs pass into code and code passes into algorithms and then algorithms start to create new algorithms, it gets farther and farther from human agency. Software is released into a code universe which no one can fully understand. (Ullman, as cited in Smith, 2018, p. 1)*

This increasing distance from the 'social' (or human design or interaction) is what we observe happening with the AutoML Zero program, where the machine learning models created are

[5] We thank a Reviewer for this point.

completely novel[6]. Though you could argue that it still contains social elements, since the original AutoML Zero program was created by humans. However, when those novel machine learning models are used to develop new machine learning algorithms, the social becomes an increasingly smaller part of it. As a consequence, with AI creating AI in rapidly repeating cycles (second, third or more-level AI) with the objective to develop better and better machine learning models (Real et al., 2020), at some point, the social and the material will no longer be inseparably entangled nor will they have to become entangled by means of actions as, eventually, the social may be so distant that it could be considered not involved at all.

While Orlikowski and Robey (1991, p. 147) earlier argued that social actions 'always involve interactions between humans', this is also no longer true; human actors now also interact with artificial social agents such as 'chatbots', some of which are intelligent. Increasingly, nonsocial actions (i.e. those performed by a 'chatbot') are beginning to look more and more like social actions. In 2020, Facebook developed the chatbot *Blender*, a chatbot that uses 9.4 billion parameters and is capable of having engaging conversations (Smith et al., 2020), though the Turing test has not yet been passed by any AI system. Further, contrary to Taylor, Groleu, Heaton and Van Every's (2001) belief that machine artifacts do not have inherent intentionality, when AI is combined with smart contracts and distributed ledger technology (i.e. blockchain), it becomes possible to develop an organisation without management or employees, but that is completely constructed using code. The result is a decentralised autonomous organisation (DAO), in which AI agents act completely autonomously and with intentionality (Van Rijmenam & Ryan, 2019).

[6] In the case of AutoML Zero, repeating cycles aim to develop better and better learning algorithms. Two or more models are randomly selected and compete against each other. The most accurate model becomes the Parent model, which clones itself into a Child model, which then gets randomly mutated. The mutated Child AI is then evaluated and paired against another model. With improved hardware and increased computing power in the coming years, it is likely that fundamentally new algorithms will be discovered, with very little to no social involved in it (Real, Liang, So & Le, 2020).

While we feel sociomateriality provides a rich theoretical lens in relation to understanding (Child) AI agency, it is also being challenged by it. We also readily observed challenges to several other key organisational theories such as actor-network theory (ANT), institutional theory and the behavioural theory of the firm. In what follows, we make initial ruminations about how each of these other theoretical lenses may make sense of AI agency and importantly, be challenged by it.

AI Agency and Actor-Network Theory

Though ANT is relatively -and unfortunately- unfashionable in current organisational theorizing, we find it important to consider when it comes to AI agents and agency. As Tatnall (2005) summarises, ANT helps to explain how agents interact with each other and allows an analysis of both artificial and non-artificial agents in the same context, avoiding the need to think in human/non-human barriers and ignoring the hierarchical distribution of actors (Latour, 2005). Such a flat ontology, where actors of different size and type are considered equally capable of creating interactions with each other, is especially relevant in conceptualizing and accounting for AI agents, and directing scholarly attention to set of relations that enable AI agents to be brought into being and through which it accomplishes its effects in everyday life (Neyland, 2019). For example, in theorizing the connections and collaborations in human-machine networks involving AI agency where the identity and source of interactions between humans and machines are increasingly indistinguishable, as is often the case even in mundane uses of AI such as 'chatbots' in call centres. When theorizing AI from an ANT perspective, human-machine networks can be viewed as nodes that share multiple, new, dimensions with each other, whereas actors not only resemble flows that interact with one another but also change those flows (Latour, 2005; Neyland, 2019), i.e. AI can change interactions among actors involved in ways yet unknown to us, which indicates an interesting new research stream.

In terms of understanding (artificial) actorhood, Latour argues an actor is semiotic: it is 'something that acts or to which activity is granted by others' (1996, p. 7), consistent with our definition of AI agency. AI agents within such human-machine networks create outputs that neither a pure social network nor a pure machine network can create independently. The behaviour of (both AI and human) actors changes the state of the network, which is itself the product of previous actions (Callon, 1990). Consequently, a self-reinforcing feedback loop can produce unexpected behaviour (Bostrom, 2014). While traditional information technologies lack reflexivity and are subject to users' whims (Leonardi, 2013) AI agents are able to change behaviour and make decisions independently and autonomously, based on previous behaviour (i.e. data) similar to how organisational decisions derived from previous behaviour shape future adoptions (Poole & DeSanctis, 1992; DeSanctis & Poole, 1994), changing context without being subject to the whims of human action. This may lead to the initial human intention behind the design of a given AI being overcome or circumvented, especially when AI starts to create its own AI. Novel programs such as AutoML Zero offer a glimpse into the future of AI as the program changes the behaviour of the machine learning model and then examines how this influences the new program, adapting it for a future iteration accordingly, which seems contrary to Archer's (2007) argument that AI does not possess reflexivity. For us this raises important questions for theorizing AI agency through an ANT lens: How can an ANT approach assist in theorizing the intentionality of AI agents? Is Callon's (1984) notion of *interessement*—actions taken by some actors to impose upon and stabilise other actors—sufficient to theorise how AI agents can be prevented from behaving in ways that do not comply with developers' original intentions or should we revise the idea of a flat ontology since AI is no longer equally capable of creating interactions, but is superior to human actors? Moreover, how can ANT theorise the interactions between humans and intelligent, autonomous, machines when ANT does not distinguish between human action and the behaviour of things (Collins & Yearley,

1992) and fails to address what role human's morality and convictions play when these two interact (McLean & Hassard, 2004)?

AI Agency and Institutional Theory

When we consider institutional theory (currently a dominant approach in organisational theorizing), digitally enabled institutional arrangements such as new organisational forms are increasingly causing changes in many industries and fields (Hinings et al., 2018). There are several opportunities for institutional theorists to theorise AI agency, including as an actor, as a mechanism that increases the speed, magnitude or direction of institutionalisation or deinstitutionalisation, as a form of institutional infrastructure, and as a diffusion mechanism.

The term 'actor' has become 'one of the central, if not most frequently evoked constructs in institutional theory', and yet its 'specification and use are contested' (Hwang & Colyvas, 2019, p. 2). In developing an abstraction of the term to assist in developing actor-based approaches to institutional theorizing, Hwang and Colyvas (2019) posit 'actor' as consisting of three core elements: '1) the level of society that claims about actors inhabit (e.g. from individual to organisational to societal); 2) the degree of generality that claims about actors occupy (i.e. from more concrete to more abstract); and 3) the ontology, meaning the essential features of an actor that determine the inclusion of social entities into the construct' (p. 5). As such, immediate questions arise for institutional theorists examining AI agency and actorhood such as: what 'level of society' do AI actors inhabit? What are the theoretically relevant features of AI actors that will provide cognitive adequacy and generalisability across the many empirical contexts in which we observe AI?

AI agency also complicates or perhaps extends ideas of institutional processes such as institutionalisation and deinstitutionalisation. Processes of institutionalisation render practices, forms, ideas and meanings taken-for-granted. What role may AI agents now play in this process, in rapidly and perhaps less obviously entrenching institutionalizing certain practices, further institutionalizing bias or inequality (Eubanks, 2018; Amis et al., 2019)? Alternatively, what is there role in deinstitutionalisation, given we are observing globally the technological challenges to many institutions underpinning democracy (elections, media) (Lindebaum et al., 2019). What institutional conditions are at stake when AI agency is introduced? How does it shape the direction, pace and sequencing of change?

Additionally, a more recent stream of theorizing on institutional infrastructure (Greenwood et al., 2011; Hinings, et al., 2017; Zietsma et al., 2017; Logue & Grimes, 2019) may also inform how to conceptualise AI agency. Institutional infrastructure is described as the 'cultural, structural and relational elements that generate the normative, cognitive and regulative forces that reinforce field governance' (Hinings et al., 2017, p. 163). These elements give rise to and maintain the stability of the social environment—the normative, cognitive, and regulative factors that specify, for instance, how organisations should interact and exchange. Such infrastructure varies in quality—specifically, in its degree of elaboration. In this regard, could AI agents provide a new form of relational infrastructure in fields? Is the work performed by AI agents more structure than agency, and when and how might this matter for institutional theorizing? And while field boundaries may be created or reinforced by the activities and role of AI agents, how might the same agents dilute or deinstitutionalisation field boundaries or professional jurisdictions? For example, as we are seeing in AI agents providing professional health or financial (robo)advice, challenging professional jurisdictions in health, finance and law. How does AI agency change understandings of negotiation processes within fields if AI agents can make their own decisions independent of humans, and the interaction and mutual dependence between and across fields and subfields (Furnari, 2016)? In this regard, AI agency may also by conceptualised as a diffusion mechanism, a carrier of certain ideas and values, that through the workings of AI agent lead to their amplification.

AI Agency and the Behavioural Theory of the Firm

Finally, we consider how the behavioural theory of the firm might be tested by AI agency. Cyert and March's (1963) work challenged neoclassical economic assumptions about the firm, introducing notions of uncertainty, conflict, satisficing behaviour and bounded rationality into explanations of firm processes, decision-making, actions and, ultimately, behaviour (March & Simon, 1958; Gavetti et al., 2012). It catalysed further research on organisational learning and the cognitive foundations of firm strategy (Levitt & March, 1988). However, AI agents seemingly overcome these previously identified weaknesses in economic models of the firm, challenging many premises of the behavioural theory of the firm (Baum & Haveman, 2020). For example, as compared to human agents, AI agents can be decidedly rational and unrelenting by design to maximise an objective and not satisfice; algorithms do as told while ignoring other considerations (Lindebaum et al., 2019). AI agents have the capability to automate decision-making and processes within organisations (Galbraith, 2014; Seidel, 2018; Swan, 2015; Van Rijmenam & Ryan, 2019). Consider the Hong Kong venture capital firm Deep Knowledge Ventures, who appointed an algorithm to its board of directors to improve decision-making (Hanson, 2017). With AI agents becoming more advanced, this could result in fully automated organisations, where human agents are managed by artificial agents (Curchod et al., 2019; Ma et al., 2018; Rosenblat & Stark, 2016; van Rijmenam, 2019). Already, we can see the first consequences of what this means within Uber, where AI agents control human drivers (Rosenblat & Stark, 2016).

AI agents, as shown in our earlier examples of strategic game-playing, have vast and superior 'search' capabilities; they are able to process vast amounts of knowledge nearly instantly and analyse consequences, thereby challenging or perhaps changing the meaning of bounded rationality (from cognitive limitations to structural limitations by AI developers/designers). Do AI agents have 'expectations' that govern decisions such as when a search process should stop? How do AI agents make decisions around resolving conflict and avoiding uncertainty? We begin to see such questions exemplified by empirical developments such as DAOs. DAOs are complex mechanisms that operate autonomously, automatically conform to compliance (Swan, 2015) and radically change decision-making within organisations (Ziolkowski et al., 2018), due to a combination of distributed ledger technologies and AI. In the coming years, and as evidenced by the AutoML Zero program and the AI breakthroughs of Deepmind and OpenAI, AI will become increasingly advanced and detached from the social, potentially behaving differently than humans. When that happens, it seems that AI potentially requires a revised science of the organisation.

Conclusion and Research Agenda: AI Agency and a Revised 'Science of the Organisation'

Rapid developments within the field of artificial intelligence increasingly results in autonomous AI agents displaying reflexivity that can act with intentionality. When AI creates AI, it is increasingly further removed from human design or interaction. This wave of technological development, we argue, is substantively different from others in fundamentally changing the human-technology interface and the cyber-physical systems in which we organise. And yet, similar to other waves of technological developments, also challenges organisational theorizing. In this essay, we argue that this looming form of AI challenges our assumptions of agency, structure, materiality, actorhood and intentionality across many perspectives of organisational, management and innovation theorizing.

To make our case we have shown how AI agency is made sense of by different theoretical perspectives, and yet also challenges the assumptions and core concepts in many of these same theories. In practice, AI agents change the nature of organisational design, decision-making, strategy, knowledge production and learning, power and governance. Beyond our observations on how this may challenge several core and relevant theories, this has further theoretical implications for leadership and contingency theories that view organisations (Curchod et al., 2019; Donaldson, 2008; Fiedler et al., 1998) in needing to accommodate AI leaders with different objectives, rules and norms than human leaders. Additionally, conceptualizing AI as an independent agent within organisations has implications for governance theories especially agency theory, as traditional governance practices involved with human actors no longer apply to AI agents (Someh et al., 2016; Bostrom, 2014; Van Rijmenam & Schweitzer, 2018).

The rapid advancements in AI and the development of AI creating AI has compelled us to take these initial steps in examining how our organisational theories may be challenged and changed by these technological developments. In each theoretical perspective outlined, we have provided a set of further theoretical questions that lay the groundwork for a research agenda that goes beyond the empirics of AI investigations, to theorizing a potential new science of the organisation. In so doing, this essay presents a call for theoretical reflexivity on the part of organisational scholars examining AI agency and actorhood.

References

Abadi, M., & Andersen, D. G. (2016). Learning to protect communications with adversarial neural cryptography. *arXiv preprint arXiv:1610.06918*.

Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy, 47*, 1–6.

Amis, J. M., Munir, K. A., Lawrence, T. B., Hirsch, P., & McGahan, A. (2018). Inequality, institutions and organisations. *Organisation Studies*, *39*(9), 1131-1152.

Armstrong S. (2017). Good and safe uses of AI Oracles. *ArXiv171105541* Cs

Armstrong, S., Sandberg, A., & Bostrom, N. (2012). Thinking inside the box: Controlling and using an oracle ai. *Minds and Machines, 22*(4), 299–324.

Bailey, D., Faraj, S., Hinds, P., von Krogh, G., & Leonardi, P. (Eds.). (2019). Special Issue of *Organisation Science*: Emerging technologies and organizing. *Organisation Science, 30*(3), 642–646.

Baum, J.A. and Haveman, H.A., 2020. *Editors' Comments: The Future of Organisational Theory*, Academy of Management Review, Vol. 45, No. 2, 268–272.

Bell, G. (2017). 2017 Boyer Lectures: Fast, smart and connected: What is it to be Human, and Australian, in a digital world [radio series]. In S. Spark (Producer), *2017 Boyer Lectures*. ABC Radio National. From https://www.abc.net.au/radionational/programs/boyerlectures/series/2017-boyer-lectures/8869370

Bell, G. (2018). Making life: A brief history of human-robot interaction. *Consumption Markets & Culture, 21*(1), 22–41.

Benaich, N., & Hogarth, I. (2019). *State of AI 2019*. From https://www.stateof.ai/

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford, UK: OUP.

Bostrom, R., & Heinen, S. (1977, December). MIS problems and failures: A socio-technical perspective, part II: The application of socio-technical theory. *MIS Quarterly, 1*(4), 11–28.

Callon, M. (1984). Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *The Sociological Review, 32*(S1), 196–233.

Callon, M. (1990). Techno-economic networks and irreversibility. *The Sociological Review, 38*(S1), 132–161.

Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-machine Systems*, *11*(4), 190–202.

Carlile, P. R., Nicolini, D., Langley, A., & Tsoukas, H. (2013). *How matter matters: Objects, artifacts, and materiality in organisation studies*. Oxford UK: OUP.

Collins, H. and Yearley, S. (1992). 'Epistemological chicken'. *In Pickering, A. (Ed.), Science, Practice and Culture*. Chicago, IL: University of Chicago Press.

Corea, F. (2019). AI knowledge map: How to classify AI technologies. In *An introduction to data: Everything you need to know about AI, big data and data science* (pp. 25–29). Switzerland: Springer.

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2019). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*. DOI: 10.1177/0001839219867024

Cyert, R. M., & March, J. G. (1963). *A behavioural theory of the firm*. Englewood Cliffs, NJ, 2(4), 169-187.

Deepmind. (2018). *AlphaStar: Mastering the real-time strategy game StarCraft II*. From https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/

Deepmind. (2020). *Agent57: Outperforming the human Atari benchmark*. From https://deepmind.com/blog/article/Agent57-Outperforming-the-human-Atari-benchmark

DeSanctis, G., and Poole, M.S. (1994). "Capturing the complexity in advanced technology use: Adaptive structuration theory*." Organisation science* 5 (2):121-147.

Donaldson, L. (2008). *The conflict between contingency and institutional theories of organisational design Designing organisations* (pp. 3-20): Springer.

Dourish, P., & Mazmanian, M. (2011, June). Media as material: Information representations as material foundations for organisational practice. Paper presented at the *Third International Symposium on Process Organisation Studies*, Corfu, Greece.

Dove, G., & Fayard, A. L. (2020, April). Monsters, Metaphors, and Machine Learning. *In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-17).

Emirbayer, M. and Mische, A., 1998. What is agency?. *American journal of sociology*, 103(4), pp.962-1023.

Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. Management sciences, models and Techniques. In C. W. Churchman & M. Verhulst (Eds.), *Management science: Models and techniques* (Vol. II) (pp. 83–97). London, UK: Pergamon Press.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

Faraj, S., Pachidi, S. and Sayegh, K., 2018. Working and organizing in the age of the learning algorithm. *Information and Organisation*, 28(1), pp.62-70.

Faulkner, P., & Runde, J. (2012). On sociomateriality. In P. Leonardi, B. Nardi, & J. Kallinikos (Eds.), *Materiality and organizing: Social interaction in a technological world* (pp. 49–66). Oxford, UK: OUP.

Fiedler, F. E. (1964). *A contingency model of leadership effectiveness. Advances in experimental social psychology*, 1(1), 149-190.

c, J., 2019. Theorizing the Digital Object. *MIS Quarterly*, 43(4).

Fleming, P. (2019). Robots and organisation studies: Why robots might not want to steal your job. *Organisation Studies*, *40*(1), 23–38. https://doi.org/10.1177/0170840618765568

Flyverbom, M. (2019). *The Digital Prism*. Cambridge University Press.

Gartner (2019). Gartner Survey Shows 37 Percent of Organisations Have Implemented AI in Some Form. From https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organisations-have

Gavetti, G., Greve, H. R., Levinthal, D. A., & Ocasio, W. (2012). The behavioural theory of the firm: Assessment and prospects. *The Academy of Management Annals*, *6*(1), 1–40.

Gill, Karamjit S. 2016. Artificial super intelligence: beyond rhetoric. *AI & SOCIETY* 31 (2):137-143. doi: 10.1007/s00146-016-0651-x.

Glisson, C. A. (1978). Dependence of technological routinisation on structural variables in human service organisations. *Administrative Science Quarterly*, 383–395.

Goertzel, Ben, and Cassio Pennachin. 2007. *Artificial general intelligence*. Vol. 2: Springer.

Hassan, N. R. (2016). A brief history of the material in sociomateriality. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, *47*(4), 10–22.

Haenlein, M. and Kaplan, A., 2019. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review*, 61(4), pp.5-14.

Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent? *Ethics and Information Technology*, *11*(1), 19–29.

Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. *Information and Organisation, 28*(1), 52–61.
Hinings, C., Logue, D., & Zietsma, C. (2017). Fields, institutional infrastructure and governance. In R. Greenwood, C. Oliver, T. Lawrence & R. Meyer (Eds.), *The Sage handbook of organisational institutionalism* (pp. 170–197). UK: SAGE.
Huber, G. P. (1990). A theory of the effects of advanced information technologies on organisational design, intelligence, and decision making. *Academy of Management Review*, *15*(1), 221–254.
Hwang, H., & Colyvas, J. (2019). Ontology, levels of society, and degrees of generality: Theorizing actors as abstractions in institutional theory. *Academy of Management Review*. https://doi.org/10.5465/amr.2014.0266
Johnson, D.G. and Verdicchio, M., 2019. AI, agency and responsibility: the VW fraud case and beyond. *AI & SOCIETY*, 34(3), pp.639-647.
Kallinikos, J. (2012). Form, function, and matter: Crossing the border of materiality. In P. Leonardi, B. Nardi, & J. Kallinikos (Eds.), *Materiality and organizing: Social interaction in a technological world* (pp. 67–87). Oxford, UK: OUP.
Kandel, A. (1991). *Fuzzy expert systems*. London, UK: CRC Press.
Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410.
Latour, B. (1996). On actor-network theory: A few clarifications. *Soziale Welt, 47*(4), 369–381.
Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford, UK: OUP.
Le, Q., & Zoph, B. (2017). Using machine learning to explore neural network architecture. From https://research.googleblog.com/2017/05/using-machine-learning-to-explore.html
LeCun, Y. (2020). *Self-Supervised Learning.* Keynote presented at the AAAI 2020, retrieved from https://www.youtube.com/watch?v=UX8OubxsY8w
Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Boston, NY: Houghton Mifflin Harcourt.
Leonardi, P. M. (2013). Theoretical foundations for the study of sociomateriality. *Information and Organisation, 23*(2), 59–76.
Leonardi, P. M., & Barley, S. R. (2010). What's under construction here? Social action, materiality, and power in constructivist studies of technology and organizing. *Academy of Management Annals, 4*(1), 1–51.
Leonardi, P. M., Nardi, B. A., & Kallinikos, J. (Eds.). (2012). *Materiality and organizing: Social interaction in a technological world*. Oxford, UK: OUP.
Levitt, B., & March, J. G. (1988). Organisational learning. *Annual Review of Sociology, 14*(1), 319–338.
Lewis, M., Yarats, D., Dauphin, Y., Parikh, D., & Batra, D. (2017, 14 June). *Deal or no deal? Training AI bots to negotiate*. From https://code.facebook.com/posts/1686672014972296/deal-or-no-deal-training-ai-bots-to-negotiate
Logue, D., & Grimes, M. (2019). Platforms for the people: Enabling civic crowdfunding through the cultivation of institutional infrastructure. *Strategic Management Journal*. https://doi.org/10.1002/smj.3110
Luca, M., Kleinberg, J. O. N., & Mullainathan, S. (2016). Algorithms need managers, too. *Harvard Business Review, 94*(1), 96–101.

Mazmanian, M., Cohn, M. L., & Dourish, P. (2014). Dynamic reconfiguration in planetary exploration: A sociomaterial ethnography. *MIS Quarterly, 38*(3), 831–848.
Missinato, A. (2018). *Deep coding: When the machine learns to code by itself*. From https://www.spindox.it/en/blog/deep-coding-ai-2/
McLean, C., & Hassard, J. (2004). Symmetrical absence/symmetrical absurdity: Critical notes on the production of actor-network accounts. *Journal of Management Studies*, 41, 493-519.
Morgan, G., & London, S. (1998). Images of organisations. *Human Resource Management Journal*, 8(2), 93.
Muñoz, F. F., & Encinar, M. I. (2014). Agents intentionality, capabilities and the performance of systems of innovation. *Innovation: Organisation & Management*, 16(1), 71-81. doi:10.5172/impp.2014.16.1.71
Murray, A; Rhymer, J., Sirmon, D (2020) Humans and technology: forms of conjoined agency in organisations. *Academy of Management Review* (in press)
Neyland, D. (2019). *The Everyday Life of an Algorithm*. Palgrave Pivot, Cham.
Noorman M, Johnson DG (2014) Negotiating autonomy and responsibility in military robots. *Ethics Inf Technol* 16(1):51–62
Norman, D. A., & Draper, S. W. (1986). *User centered system design: New perspectives on human-computer interaction*. Boca Raton, FL: CRC Press.
Omohundro S (2016) Autonomous technology and the greater human good. In: *Müller V (ed) Risks of artificial intelligence*. CRC Press, Boca Raton, pp 9–27
OpenAI. (2019). *OpenAI Five*. From https://openai.com/five/
O'Neil, Cathy. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group (NY).
Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organisations. *Organisation science*, 11(4), 404-428.
Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organisation Studies, 28*(9), 1435–1448.
Orlikowski, W. J. (2009). The sociomateriality of organisational life: Considering technology in management research. *Cambridge Journal of Economics, 34*(1), 125–141.
Orlikowski, W. J., & Robey, D. (1991). Information technology and the structuring of organisations. *Information Systems Research, 2*(2), 143–169.
Orlikowski, W. J., & Scott, S. V. (2008). 10 sociomateriality: Challenging the separation of technology, work and organisation. *Academy of Management Annals, 2*(1), 433–474.
Perrow, C. (1967). A framework for the comparative analysis of organisations. *American Sociological Review, 32*(2), 194–208.
Poole, M. S., G. DeSanctis. (1992). Microlevel structuration in computer-supported group decision making. *Human Comm. Res.* 19(1) 5–49.
Radford, A., Jeffrey, W., Child, R., Luan, D., Dario, A., & Sutskever, I. (2019). *Language models are unsupervised multitask learners*. From https://openai.com/blog/better-language-models/
Rao, A., & Verweij, G. (2017). *Sizing the prise: What's the real value of AI for your business and how can you capitalise?* From https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prise-report.pdf
Real, E., Liang, C., So, D.R. and Le, Q.V., 2020. *AutoML-Zero: Evolving Machine Learning Algorithms From Scratch*. arXiv preprint arXiv:2003.03384.
Riemer, K. and Johnston, R.,. (2017). Clarifying ontological inseparability with heidegger's analysis of equipment. *MIS Q.* 41, 4 (December 2017), 1059–1081.

Rosenblat, A. and Stark, L., 2016. Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication*, 10, p.27.
Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Malaysia: Pearson Education Limited.
Scott, S. V., & Orlikowski, W. J. (2012). Reconfiguring relations of accountability: Materialisation of social media in the travel sector. *Accounting, Organisations and Society*, *37*(1), 26–40.
Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organisational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66–83.
Silver, D. and Hassabis, D., 2017. *AlphaGo Zero: Starting from scratch*. London, United Kingdom: Deepmind.
Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.
Smith, A. (2018, 30 August). Franken-algorithms: The deadly consequences of unpredictable code. *The Guardian*. From https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger
Smith, E.M., Williamson, M., Shuster, K., Weston, J. and Boureau, Y.L., 2020. *Can You Put it All Together: Evaluating Conversational Agents' Ability to Blend Skills*. arXiv preprint arXiv:2004.08449.
Someh, I.A., Breidbach, C.F., Davern, M.J. and Shanks, G.G., 2016, June. *Ethical Implications of Big Data Analytics*. In ECIS (pp. Research-in).
Sørensen, M.H. and Ziemke, T., 2007. Agents without agency?. *Cognitive Semiotics*, 1, pp.102-124.
Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge, UK: Cambridge University Press.
Swan, M. (2015). *Blockchain: Blueprint for a new economy*. Sebastopol, CA: O'Reilly Media.
Tatnall, A. (2005). Actor-network theory in information systems research, in *Encyclopedia of Information Science and Technology*, First Edition. IGI Global. p. 42-46.
Taylor, J. R., Groleu, C., Heaton, L., & Van Every, E. (2001). *The computerisation of work: A communication perspective*. Thousand Oaks, CA: SAGE.
Tseng, C.-Y., & Ting, P.-H. (2013). Patent analysis for technology development of artificial intelligence: A country-level comparative study. *Innovation: Organisation & Management*, 15(4), 463-475. doi:10.5172/impp.2013.15.4.463
Turchin, A. and Denkenberger, D., 2020. Classification of global catastrophic risks connected with artificial intelligence. *AI & SOCIETY*, 35(1), pp.147-163.
Van Lent, M., Fisher, W., & Mancuso, M. (2004, July). An explainable artificial intelligence system for small-unit tactical behaviour. Paper presented at the *19th National Conference on Artificial Intelligence*, San Jose, California, US.
Van Rijmenam, M. (2019). *The organisation of tomorrow: How AI, blockchain and analytics turn your business into a data organisation* (Vol. 1). London, UK: Routledge.
Van Rijmenam, M., & Ryan, P. (2019). *Blockchain: Transforming your business and our world* (1st ed.). London, UK: Routledge.
von Krogh, G. (2018). Artificial intelligence in organisations: New opportunities for phenomenon-based theorizing. *Academy of Management Discoveries, 4*(4), 404–409.
Wace Peck, M. (2011). *Internet 'has same weight as a strawberry'*. From http://www.digitaljournal.com/article/311046

Wang, F.-Y., Zhang, J. J., Zheng, X., Wang, X., Yuan, Y., Dai, X., … Yang, L. (2016). Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA Journal of Automatica Sinica, 3*(2), 113–120.

Weißenfels, S., Ebner, K., Dittes, S., & Smolnik, S. (2016, January). Does the IS Artifact matter in sociomateriality research? A literature review of empirical studies. Paper presented at the *49th Hawaii International Conference on System Sciences*, Hawaii, US.

World Intellectual Property Organisation (2019). *WIPO Technology Trends 2019: Artificial Intelligence.* World Intellectual Property Organisation.

Woodward, J. (1965). *Industrial organisation: Theory and practice* (Vol. 3). London, UK: OUP.

Yang, Qiang, Yu Zhang, Wenyuan Dai, and Sinno Jialin Pan. *Transfer learning*. Cambridge University Press, 2020.

Yudkowsky, E. (2007). *Levels of organisation in general intelligence*, in Artificial general intelligence. Springer. p. 389-501.

Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. *Global catastrophic risks*. 1: p. 303.

Zammuto, R. F., Griffith, T. L., Majchrzak, A., Dougherty, D. J., & Faraj, S. (2007). Information technology and the changing fabric of organisation. *Organisation Science, 18*(5), 749–762.

Ziolkowski, R., Miscione, G., & Schwabe, G. (2018). Consensus through blockchains: Exploring governance across inter-organisational settings. In Association for Information Systems (Ed.), *Proceedings of the International Conference on Information Systems (ICIS)*. From https://aisel.aisnet.org/icis2018/governance/Presentations/10/

Zittrain, J. L. (2006). The generative internet. *Harvard Law Review, 119*(7), 1974–2040.