The Emerald Handbook of Authentic Leadership

Chapter 6

The Search for Authenticity in Artificial-Intelligence-Enhanced Leadership

Niilo Noponen, Tommi Auvinen and Pasi Sajasalo

Abstract

This chapter critically examines whether it may be possible to create an AI-based authentic leader, questioning the inherent contradiction between artificial and authentic. The authors pose central research questions: Does the application of AI – even just as a powerful resource – challenge the tenets of authentic leadership? What are the possibilities and limitations of the concept of authenticity in AI-based management systems? Moreover, with the help of three vignettes illustrating practical applications of AI-based systems in leadership and management tasks, the authors illustrate how technology may be used to either control or empower workers and leaders. The authors call for research to assess whether the search for authenticity in AI-based leadership could lead anywhere, warning that it could entrap us in unresolvable existential and conceptual ambiguity, ultimately diverting our focus from the essence of leadership altogether.

Keywords: Authentic leadership; authenticity; artificial intelligence; algorithmic management; digitalisation; avatar; future of work

Introduction

In an experiment to find out whether artificial intelligence (AI) has potential as an expert informant in decision-making, the Committee for the Future of the Parliament of Finland interrogated a GPT-3-based Project December AI, 'Samantha', on how to solve global poverty. Samantha's solution to eradicating global poverty was startling: 'We must intimidate them [the rich]. I mean, we must take the Parliament hostage. If they do not listen to us, we must kill some of them.' (Iltalehti, 19 April 2021) This is an example of how contemporary leaders may seek assistance from AI: as a decision-making support system. While the AI-produced solution appears extreme, it is certainly not a foreign idea relative to those of radical human political thinkers. In fact, renowned Finnish film director Aki Kaurismäki proposed a nearly identical solution to the same problem a decade prior: 'The only way for mankind to get out of this misery is to kill the 1% who own everything [...] The rich. And the politicians who are the puppies of the rich' (The Guardian 4 April 2012).

One could argue that the AI merely emulates the 'prevailing reality' based on the sources used to train it. The solution offered by Samantha may resonate authentically with the perspectives of human actors. Whether the AI solution is socially or morally acceptable – and the same goes for identical views presented by human actors – is another question.

While the example may be an extreme case of AI as an 'expert advisor' for decision-making, cyborgs, androids and other technological creatures imitating human forms or behaviour have fascinated human minds for centuries. While such smart machines call to mind images of the future, the most interesting aspect of them is not how they are depicted in fiction but how they reflect humanity The comparison of humans and human-like machines resonates with

the unique aspects of what makes us what we are. Are ‘authentic’ and ‘artificial’ fundamentally contradictory concepts? The idea of AI – a machine behaving in a way that would be considered intelligent (McCarthy et al. 1955) and simultaneously appearing authentic – could present an interesting dynamic to the concept of authentic leadership.

The hype around AI and the capacity of AI, including applications to leadership and especially management (Harms and Han 2019), seems infinite. Studies (e.g., Lee et al. 2015; Rosenblat and Stark 2016) demonstrated that first-wave algorithmic management systems are often disliked by workers. Is this distaste merely due to novelty, or does it have to do with a perceived lack of authenticity? Could technology conceivably provide an ‘authentic artificial leader’, or does the contradiction between authentic and artificial make such a leader impossible? What does ‘authenticity’ mean in this context?

In this chapter, we focus on the possibilities and limitations of advanced digital technologies, such as AI-based management systems, in the context of authentic leadership and explore perspectives of the authenticity of AI applied in a leadership context. To evaluate the extent to which an artificial leader entity based on AI may appear authentic, we approach leadership from a behavioural perspective (Hannah et al. 2014). Using contemporary examples, we construct three vignettes to illustrate the implications of AI for leadership authenticity.

Authenticity, Authentic Leadership and AI – Like Oil and Water?

In the exploration of authenticity, existentialist philosophers’ offer important insights. According to Sartrean thought, existence precedes essence, meaning a person is free to choose their beliefs and desires. In these terms, authenticity may be understood as being true

to the originality of one's being (Holt 2012) or the degree to which a person's actions are aligned with their beliefs and desires (e.g., Harter 2002, 382).

The relationship between technology and authenticity is not necessarily harmonious. Heidegger believed that while the making and producing of 'artificiality' is a universal human condition, modern technology reduces us to clever animals with no insight into our own authenticity (Zimmerman 1990, 221). This notion implies that there is a natural way of being that may be obstructed as we use technology to cope with our existence. It is easy to see this conflict in the way social media platforms polarise user behaviour.

While philosophical discussion of authenticity has historic roots stretching back to Socrates (see Nehamas 1999), discussion of authentic leadership theory emerged at the turn of the millennium to answer a call for a new type of leadership amidst high-profile cases of unethical behaviour by political and business leaders (Iszatt-White and Kempster 2019). Building on positive psychology, authenticity in authentic leadership theorising rests on four 'pillars' (Kempster et al. 2018, 320), as outlined by Avolio and Gardner (2005): self-awareness, balanced processing, relational transparency and authentic behaviour/action. Regardless of high aspirations, authentic leadership theory has not saved us from the unethical behaviour of people in power. This being the case, would removing the human factor lead to something we could call truly authentic leadership?

During the last decade, new forms of leadership/management, previously seen only in science fiction, have emerged as technology has advanced. Organisations from Uber to Amazon have resorted to AI-based technologies to supplant leaders by concentrating on efficiency and control over people in the spirit of Taylorism, which has resulted in intense monitoring (e.g.,

Meyer 2016, ix-xii) and shifting monitoring to AI so it can be performed automatically (e.g., Parry et al. 2016; Höddinghaus et al. 2021). This has opened an avenue – or possibly Pandora's Box – for organisations to lead/manage their workforce with little or no human involvement.

For those frustrated by unethical, corrupt leaders, this may seem like discovering the philosopher's stone of leadership. Imagine a system able to emulate the best qualities of leaders while discarding the undesirable ones. After all, leadership scholars have long been on a quest to unveil the qualities of a 'good leader' in leadership theory building (e.g., Ciulla 2018). In revealing such qualities, researchers have tried to identify the ideal traits, features and practices, or 'authentic' behaviour, of an effective leader to promote social processes in achieving organisational goals. However, Hannah et al. (2014) criticise the tendency of leadership research to entangle types of leadership and leaders and disregard the behavioural orientation of many of the 'newer' leadership theories, such as authentic leadership:

> [R]esearchers … too often write theory to support hypotheses using language that tends to describe "types" of leaders (i.e., who transformational [or authentic, participative, etc.] leaders "are" or what they "possess") versus types of behaviour they enact (i.e., what leader actions are transformational [or authentic, participative, etc.]). There are meaningful differences in stating "transformational leaders are charismatic individuals who…" versus stating "when leaders act in ways perceived by followers as charismatic…". The former tends to anthropomorphize a "super leader," whereas the latter focuses on ways a leader can act and be perceived by followers. (Hannah et al. 2014, 600)

Despite the continuing search for super leaders, even the most influential or charismatic leaders – be they Mahatma Gandhi, Martin Luther King Jr., Nelson Mandela, Barack Obama or Donald J. Trump – are authentic precisely because of human qualities related to intelligence, personality, temperament and judgment, including shortcomings in their conduct, as leaders or otherwise. Leaders are authentic if they remain true to themselves and because others – their followers – attribute this status to them because of their behaviour. Authenticity cannot be 'declared' but only 'earned', as all leadership is relational (e.g., Avolio and Gardner 2005). At the same time, 'faultless' AI systems optimised with millions of simulations might appear an ideal solution to lead organisations and address global problems – similarly to Samantha operating with impeccable logic above – that authentic leadership sets out to solve. Would authenticity necessitate human qualities, including shortcomings such as 'flawed' judgment, to allow someone to exist as their unique self or be true to the originality of their being, as argued by Holt (2012)?

Designating the status of being true to one's own being or existing as one's 'true' unique self (i.e., establishing authenticity) is beyond an on/off dichotomy. As pointed out by Avolio and Gardner (2005), no person is ever perfectly authentic or inauthentic, but there is a scale of authenticity that one reaches at various points and times. If it is difficult to evaluate the authenticity of a flesh-and-blood leader, the same applies to an artificially intelligent one, as it is far from clear by which standards authenticity should be judged.

Machines and computer software, regardless of how 'intelligent' they may appear, do not have beliefs and desires the way humans do. We can command software algorithms to perform towards the desired goal and establish the rules and conditions they are to follow and arguably determine what could be termed the 'beliefs and desires' of AI systems. Doing so

presents a myriad of challenges, such as AI adopting the biases of its designers or the data on which it is trained, which could produce unintended outcomes. Anecdotal examples include Microsoft's chatbot Tay becoming a foul-mouthed racist almost overnight, Amazon's AI recruiting tool turning into a misogynist and a bank's loan engine learning to discriminate against people of colour based on the applicant's address (Tech Times 2019) and disenfranchising them.

Based on the premise that an AI system's goals are similar to a human being's core beliefs, an AI system may exhibit authentic behaviour if it simply performs its tasks accurately. Accepting this premise, making an AI-based management system manifest what could be regarded as a leader's authenticity becomes solely a matter of training the software algorithms to perform the tasks of authentic leadership. However, while authenticity itself is an elusive concept, things become even more complicated when it is combined with leadership. Authentic leadership is a contested topic. Einola and Alvesson (2021) question the viability of combining the concepts of authenticity and leadership, as they consider them fundamentally opposing phenomena – the former looking inwards to know oneself, the latter seeking to exert influence outwards on others.

Organisational and societal structures have an impact on perceptions of authenticity, further complicating the issue. Authentic leadership theory has been criticised for not acknowledging that modern workplaces are seldom hospitable environments for authenticity (Alvesson and Einola 2019). A leader's authenticity is limited by the degree to which the goals and beliefs of the company are aligned with their own. If a firm's main goal is to maximise shareholder value, it is more difficult for a leader who does not believe in shareholder primacy over other stakeholder groups to act authentically. Conversely, an AI-based management system can be

programmed to carry out the organisation's goals without making value judgements based on discrepant preferences. An AI-based leader would act in accordance with company goals and thus appear highly authentic by operating according to its 'core beliefs'.

This brings us to the relationship between artificial and authentic. While viewpoints to the contrary exist (Whelchel 1986; Heidegger 1977), we believe that technology is a tool. A hammer can be used for construction or destruction, AI is not dissimilar: an AI system performs exactly the task it is taught to perform. Although AI programmes are not self-conscious, an AI-based leader could be infallibly authentic to the 'desires and beliefs' coded as its operational goals and display the behaviour that follows from them. However, any AI software is only as good as the data provided to train it. If a machine learning algorithm learns from 'faulty' human behaviour (judged against societally held criteria), it will exhibit equally faulty algorithmic behaviour, or, like Samantha, take it to the extreme. This phenomenon may present a way to uncover systematic biases or unethical behaviour. In addition, as the reinforcement model of machine learning, discussed below demonstrates, not all approaches to training algorithms are tied to human behaviour.

Having acknowledged the complexities and elusive nature of authenticity, in the following two sections we evaluate, through the lens of authentic leadership, what happens when technology assumes leadership/management functions to build a basis for our concluding discussion and implications for authentic leadership.

From Algorithmic Control to Digital Empowerment: Current AI Usage in Management

In this section, we focus on how management systems afforded by advanced digital technologies are utilised by contemporary organisations and why such systems may or may not manifest authenticity.

AI was defined by McCarthy (1955, 11) as ‘making a machine behave in ways that would be called intelligent if a human were so behaving’. Although this definition has evolved to reflect developments in the field, McCarthy accurately describes how we perceive AI in the context of leadership and management in this chapter. Any technological system able to perform the tasks of a manager or leader may be considered artificially intelligent. As such, we see AI as an umbrella term under which there have been developments in the associated technology, such as machine learning and deep learning neural networks (Jarrahi et al. 2021).

AI may seem a recent issue, but the ancient Greeks imagined artificially intelligent creatures, such as the mythical giant bronze robot Talos, 2,700 years ago (Mayor 2018). More recently, after McCarthy, Minsky and their peers coined the term in the 1950s, periods of AI excitement, development and unrealistic expectations have been followed by ‘AI winters’, during which interest and funding for further development cooled (Haenlein and Kaplan 2019). Due to the excitement–winter cycles, after the first AI system had beaten its programmer in checkers in 1959 (Samuel 1959), it took 38 years for an AI system to win against the best human chess player, in 1997.

Before the turn of the millennium, practical AI-based applications appeared, from autonomous drawing programmes to automated junk mail filters. The theoretical foundations for artificial neural networks were laid during the latter decades of the 20th century. Around 2010, enough computing power and big data became available to unleash the potential of

deep learning neural network designs (Haenlein and Kaplan 2019), resulting in several advanced AI solutions, from chatbots to self-driving cars. Vignette 6.1. illustrates some of the ways AI-based systems are used to carry out managerial/leadership tasks.

Vignette 6.1: Algorithms Managing and Controlling workers: Uber and WorkSmart by Crossover

Uber

Consider Sam. She started her career driving an Uber taxi – to be her own boss, as the company promises. Sam's only manager is a smartphone application algorithm, so she decides when, where and for how long she will work. Initially, everything goes well. Sam enjoys chatting with customers and getting to know her city better. After the first month, though, Sam has to start working longer hours because most of the jobs she gets are low-paying minimum fare rides.

As Sam drives more at night, she begins to encounter unpleasant passengers. The real trouble begins when a customer attempts to harass her. Luckily, she manages to escape, and, terrified, tries to contact the company for help and support. After 40 minutes holding the line, she finally reaches Uber's call centre in the Philippines. All she gets is a template answer on how to handle difficult situations. She tries sending emails to Uber. After a week or two, she receives a reply, promising that she will not be connected with the harassing passenger again. But what if they try to attack other drivers? Frustrated with the lack of responsibility and leadership shown by the company, Sam no longer feels engaged in the job. She is left thinking that even an algorithmic manager should be able to offer better support, enhancing employee engagement.

Sources: Lee et al. (2015); Rosenblat and Stark (2016); Ma et al. (2018); Rosenblat (2018)

WorkSmart by Crossover

The software company Crossover, specialising in recruiting remote workers for clients, takes managerial control of workers to the next level by harnessing 'technology as the master'. The company uses a surveillance system called WorkSmart to measure the productivity of remote workers. According to their website, the software 'will use keyboard activity, application usage, screenshots, and webcam photos to generate a timecard every 10 minutes'. This means that if the contracted worker fails to reach the desired level of productivity set by their contractor, they will not be paid for the 10 minutes in question. While the long-term effects of working under such

conditions of intense monitoring have not been empirically studied, Dzieza's (2020) investigative article depicts a pattern of employee exhaustion, burnout and high turnover.

Sources: 'WorkSmart Productivity Tool' (2020); Dzieza (2020)

While fictional, Sam's story is based on real-world experiences of Uber drivers, according to researchers who have examined their perceptions of working under an algorithmic management system (Rosenblat 2018). Of course, the Uber story does not represent all algorithmic management systems but illustrates one implementation. The same applies to Crossover's WorkSmart system. While Uber has automated control and surveillance, typically tasks perceived to fall within a managerial remit, the company provides few leadership services through AI-based systems/solutions or human leaders. With only transactional motivation and no social support, the drivers have had to resort to online forums for peer support and information on how to play the algorithm to their advantage. AI not only affects the workforce but leaders and managers as well: directly, by taking over tasks, and indirectly, by changing the environment in which organisations operate (Noponen 2019). As the Uber and Crossover examples show, algorithms have replaced a larger portion of managerial than leadership tasks, due to the nature of the tasks in the two areas.

There are several routine management tasks in which human involvement is rather easy to replace, especially scheduling, monitoring and even recruitment. These repetitive tasks typically produce a large amount of data, which makes training machine learning algorithms to carry them out relatively easy and allows the algorithmic management system to perform them more efficiently than a human could. Interpersonal leadership tasks such as coaching, mentoring and motivating, however, are harder to automate because of their non-routine nature and the personal consideration involved (Jarrahi 2018).

Trying to understand the implications of automated workforce management and control, scholars have studied platform economy companies (Kaine and Josserand 2019). Uber is an intriguing case as it uses a mobile phone application to control millions of drivers around the world. Similarly, Crossover's recruitment and productivity monitoring platform offers a glimpse of how control of the results of remote work may be systematised and automated. Such software algorithms assuming managerial tasks have been defined as algorithmic management (Lee et al. 2015). Many studies have begun by asking what perceptions and experiences workers under an algorithmic superior have. A summary answer is: there is considerable room for improvement in the ways Uber and similar platform economy companies have implemented algorithmic management (Rosenblat and Stark 2016; Ma et al. 2018). A more profound question is whether collaboration between leaders and followers requires authentic human interaction and 'authentic' leadership at all.

Uber's AI-powered management system aims at maximising the number of rides and resultant revenue to Uber. Apparently, it does that very capably. However, threatening situations for drivers are outside the 'core beliefs' of the system, so it does not consider such occurrences and has no way of acting upon them. Similarly, Crossover's WorkSmart AI system aims to push the productivity of the contracted worker to the maximum by unflaggingly monitoring their actions and measurable results. From the perspective of authentic leadership, both systems follow their internal logic of maximising revenue for the company by exerting managerial control with impeccable detail. Uber's and WorkSmart's systems follow the AI system's core beliefs for decision-making and action. This can be thought of as affording AI an authentic organisational actor role through behaviour that is in line with the core beliefs of the system.

Both AI systems in Vignette 6.1 may be thought of as authentic in their managerial control (i.e., behaviour), which does not necessarily differ from that expected of a highly goal-driven human actor in similar circumstances. If the goal is to squeeze every ounce of output capacity out of the workforce, pushing workers to perform by close and constant monitoring and control of work is what often ensues. However, it is not clear that increased surveillance equals increased productivity. Monitoring methods tend to create costly tasks as workers circumvent the systems or conceal their activities. In fact, some studies provide evidence that increasing privacy – implying trust-based rather than control-based relations – may improve work performance (e.g., Bernstein 2012). Thus, while AI-based management systems may perform their tasks as human actors would – virtually in real-time – and appear authentic in their behaviour, such a surveillance-oriented work environment is arguably de-humanising and portrays a human being as a cog in the machine, akin to Taylorism. The difference is that AI-enabled Taylorism is more extreme.

AI-based management systems' ability to exert extreme control heightens the role of trust between employees and employers as an issue to consider in general, but also with respect to authentic leadership (Avolio and Gardner 2005). Advanced digital technologies, such as AI, allow organisations to offer more freedom to their workforce or, conversely, resort to technology as a means of relentless monitoring and ever tighter control. While this argument may be made concerning technology-assisted control, the same may apply to flesh-and-blood managers and leaders: their role increasingly needs to transform from control towards supporting and enabling – possibly assisted by technology – for them to remain relevant to their organisational members (Schildt 2017). Vignette 6.2 introduces two companies, Buurtzorg and Vincit, to illustrate how algorithms may be employed to allow workers more self-direction and freedom. It serves as a contrast to the close control and monitoring in

Vignette 6.1 to provide an alternative perspective on the use of AI-enabled systems in organisations.

Vignette 6.2: Algorithms as enablers of shared leadership and self-leadership

Buurtzorg

The Dutch nursing company Buurtzorg is an example of technology performing a servant role. Using what they call 'hands-off management' and self-organised teams of a maximum of 11 nurses, Buurtzorg has combined high-level client and staff satisfaction with good financial outcomes. Illustrating how technology may be used to support workers, Buurtzorg uses a software platform, Buurtzorg Web, to help nurses achieve their organisational mission: holistically meeting people's needs. Even though the software tracks productivity, the system is enabling by nature, supporting self-directed teams in caregiving, communication and teamwork. Provided that they reach a baseline productivity level, each team is free to organise and make decisions as they deem most suitable. While few would classify Buurtzorg Web as an advanced AI technology, it suggests that user-friendly technology that helps employees to be efficient and productive does not need to be based on the latest developments of machine learning but, rather, bottom-up innovation in avoidance of excessive bureaucracy, administrative burden and complexity. As Nandram and Koster (2014, 181) put it, 'Trust is crucial, so control mechanisms should be limited to the team level and company-wide data should be for monitoring and benchmarking rather than control'.

Sources: Nandram and Koster (2014); Buurtzorg (2016)

Vincit

Vincit, a Finnish software company, is known for its effort to enhance employee wellbeing through leadership, having received the Best Workplace award in both Finland and Europe (The Best Place to Work in Europe 2016 – Vincit (bestworkplaceineurope.com)). In AI and leadership/human resource management (HRM), Vincit has developed a digital leadership system called Leadership as a Service (LaaS, or, more recently, 'Guidin'), which they refer to as 'a ready-made platform for modern leadership'. Instead of producing anonymous HRM statistics for leaders, Vincit flipped the system on its head by creating a genuinely employee-oriented model. Vincit LaaS started as an internal webshop for a selection of (partly automated) leadership services for the company's staff. The system allows employees to access support they need when they need it and from whom they deem fit. By using the system, employees can track and control their work and wellbeing. Instant

feedback helps Vincit to improve the service selection and focus on things that create value for their employees and the company.

Source: Vincit (2022)

From a leadership perspective, the cases in Vignette 6.2 clearly differ from those in Vignette 6.1 in their orientation. Uber's and Crossover's approach to AI-based systems (Vignette 6.1) emphasises worker control and a managerial-oriented, even authoritarian, approach, reflecting traditional hierarchical methods of control. In contrast, Buurtzorg's and Vincit's solutions in Vignette 6.2 offer employee-centred, non-hierarchical systems, allowing self-determination at work. The Vincit case, in particular, presents an approach that helps workers to identify what they expect and need from their leaders. From the leadership perspective, defining the parameters for an AI-supported leadership portal (LaaS) allows the employees to be heard and offers them agency over their work, unlike in Vignette 6.1, where the AI systems failed to do either.

The approaches to AI-based management systems in the vignettes paints two different pictures in assessing authenticity. Buurtzorg and Vincit utilise models of self-directed teams and self-leadership in which algorithms are not masters, but servants, to the organisational members. Instead of evaluating the technological systems, the issue here is about the services provided by the systems and their 'behaviour' in support of workers. The systems used by Buurtzorg and Vincit resonate with such aspects of authentic leadership as relational transparency and balanced processing (Neider and Schriesheim 2011). Buurtzorg Web is used to share information, experiences and best practices within and between teams. Similarly, an employee may use the self-training features of Vincit's LaaS system to understand their strengths and weaknesses, as if guided by a human leader, which resonates with the self-awareness aspect of authentic leadership (Neider and Schriesheim 2011).

Vignettes 6.1 and 6.2 highlight opposite approaches: technology as a master in the former and technology as a servant or a tool to support and enable workers in the latter. They emphasise the importance of trust when companies decide how to use technology and for what purpose – control or guidance. While employers may find it hard to resist the temptation to know exactly what their employees are doing, there have been warnings of the dangers of (organisational) surveillance capitalism (Zuboff 2015) and the necessity of labour laws to protect workers (Ajunwa et al. 2016). Leaders deciding whether to use such technology should be aware that the effectiveness of surveillance is questionable, as pointed out by Bernstein (2012). This would suggest that, compared to the WorkSmart tool, with its emphasis on control, the empowering, trust-based model of Buurtzorg and Vincit, in which technology plays a servant role for the organisation and users, is perceived to be fairer and enable more productivity.

Having questioned the necessity of extravagant algorithms and emphasised the importance of a horizontal organisational structure and ground-up method in developing AI-based leadership and management systems, we now shift our focus to the future to paint a picture of what the next generation of AI-based leadership systems may look like.

Training an artificial (machine) leader

How could one create an AI-based leader? In this section, we look at basic principles of machine learning and how they may apply to AI-based leadership. While there are several approaches to machine learning, most share common aspects. First, the problem to be solved needs to be defined. Next, the learning model needs training data related to the problem. To allow the model to evaluate its performance and modify its parameters after each round of

simulation, the data often need to be labelled. Finally, computing power is necessary to churn through simulations as many times as needed for the model to optimise its parameters by repeatedly testing its performance.

Unsupervised reinforcement learning is a way to train software agents – pieces of software working autonomously and continuously, functioning as agents for a user or another programme (Technopedia 2022) – without the need for labelled data. What results is a form of hyper-accelerated series of trial and error. First, the agent approaches the task at random, making questionable decisions. After a million repetitions retaining what gives correct results and discarding what gives false results, the agent will be closer to perfecting its task. This method leads to behaviour that is often different from a human expert. It has allowed computer software to win against the best human players in such games as Go and Starcraft 2 – something previously considered nearly impossible.

Computer games such as Starcraft 2 reveal the limits of AI. The game consists of playable races, with distinct skills and strategies. Although AI software can train itself to beat the world's best human player in one of the races, it would be worse than a novice if tasked to play an unfamiliar race. This is because AI commonly lacks transfer learning – the ability to use acquired skills and knowledge in new situations – a feat in which humans are relatively accomplished. For instance, previous leadership experience in a sports team may benefit a leader in an office setting. Extending the idea to the context of leadership, the rules and tactics of 'the game' of leadership may be taught to an AI-based leader, but the needs, emotions and personalities of the followers are more complicated for AI to master. Nevertheless, computer vision systems are able to recognise (yet not really understand) human emotion from speech (Lim et al. 2017) and facial expressions (Tarnowski et al. 2017),

which would be important skills for an AI-based leader. Natural language processing (NLP) allows machines to understand speech and text. An NLP-enabled AI-based leader would be able to monitor email, coffee room or office cubicle conversation at an increasingly negligible cost (Schildt 2017). However, such an AI-based leader could unintentionally bring the most questionable aspects of surveillance capitalism (Zuboff 2015) to the forefront and, in the worst-case scenario, enact an organisational version of the Orwellian dystopia. Vignette 6.3 depicts recent developments in AI technology and the possibilities NLP may offer for the construction of AI-based leaders.

Vignette 6.3: GPT-4 & Avatars by Synthesia

Generative Pre-trained Transformer

Open AI's NLP model, Generative Pre-trained Transformer (GPT-4), represents one of the latest breakthroughs and the type of technology most likely to be used in managerial tasks and leadership. With 175 billion machine learning parameters, GPT-4 can create code, columns, poetry, prose and perhaps even scientific articles. However, as evidenced by the racist and hateful dialogue it is also capable of creating, GPT-4 lacks understanding of meaning and common sense. With the rapid pace of machine learning development even the most advanced systems of today, such as GPT-4, are likely to be outdated swiftly.

Still, GPT-4 offers insight into the likely building blocks of an AI-based leader. While a single machine learning algorithm is limited to excelling in one task, it is possible to combine multiple algorithms into one system. In attempting to create an AI-based human-like leader, one could fill a virtual or robotic body with multiple technologies that enable an artificially authentic leader to see, hear and read in order to acquire information and, subsequently, create speech or writing based on the vast data acquired, allowing GPT-4 to convincingly emulate human interaction.

Synthesia

GPT-4 becomes even more impressive when the text it creates is personified by a virtual avatar, such as those generated by Synthesia. The company offers a variety of 'AI presenters' with such natural features and expressions that a viewer could easily be fooled into thinking they are watching a real person. A combination of GPT-4's powerful NLP (and decision-making)

technology embodied in a digital avatar offers a glimpse of what the artificially authentic AI-based leader could look like. We will explore the appearance-related issues of an AI system in the next section.

Figure 1 Synthesia's AI Avatar, 'Anna'. Source: Courtesy of Synthesia (www.synthesia.io).

SCAN ME

Figure 2 Video Example of an AI Avatar. Source: Courtesy of Synthesia (www.synthesia.io). Note: Use this QR code to access a video example of how an AI avatar may look and sound with Synthesia's generative AI

Sources: Brown et al. (2020); Synthesia (2022)

Future versions of a GPT-4-type NLP system could allow an AI-based leader to process information in real time and act in a manner it deems appropriate in the situation appearing

plausibly authentic visually and in behaviour. The current version provides a highly authentic first impression and is hard to distinguish from a person appearing in a video call.

Hypothetically, with enough leadership situation training data, one could build a database of proper behaviour for given situations. This is similar to the way self-driving cars or chatbots learn by encountering new situations and developing suitable patterns of action. However, communication between people in a leadership context is many times more complicated than driving in heavy traffic or renewing an insurance contract online, developing such a model represents a sizable challenge. The model would start without any experience, but with enough time, effort and computing power, the AI-based leader could reach a level where the number of completely new situations asymptotically approaches zero. Still, barring a singularity type situation where AI development takes an unprecedented leap, algorithms' inherent weakness in novel situations is likely to remain for the foreseeable future (Alkhatib and Bernstein 2019).

With a hypothetical AI-based leader building on the AI-enabled management systems described above, we move to evaluating issues related to authentic leadership in the AI context. From a historical perspective of leadership, Vignette 6.1 typifies low-qualification work, where a leader has closely supervised, rewarded and sanctioned followers according to their job performance. This represents the classical Theory X-type situation outlined by Douglas McGregor (1960). Vignette 6.2, however, typifies high-qualification expert work, or what could be called autonomous work, with a more participative and democratic leadership style, fitting what McGregor (1960) refers to as Theory Y. In both cases, it is safe to assume a human leader would act similarly and, thus, the management/leadership behaviour of the AI-based systems could be perceived as (artificially) authentic. As the virtual avatar in Vignette

6.3 shows, creating a convincing, plausibly authentic, human-like appearance for an AI system is possible.

Discussion – Artificially Authentic Leadership?

An AI system may be trained to perform well in many individual tasks of relational transparency and balanced processing (Neider and Schriesheim 2011) and, therefore, fulfil aspects of authenticity. While the black-box nature of algorithms is likely to persist, an advanced system could be taught to explain the key aspects of its decision-making to human actors. Any method of assessing authenticity, however, will contain ambiguity. The Authentic Leadership Inventory (ALI) proposed by Neider and Schriesheim (2011) is no exception. While their research supports the discriminant validity of ALI, it is unclear how ALI precisely measures authentic leadership, as distinguished from ethical or transformational leadership, or the qualities attached to good leaders in general. We have utilised the dimensions of the ALI framework (self-awareness, relational transparency, moral perspective, balanced processing) to discuss the authenticity of an AI leader in comparison with human leaders. This comparison remains ambiguous, because of the substantially different compositions of the human and machine ‘brain’ and the way these differences manifest in their behaviour.

It has been argued that the introduction of AI technologies may accelerate horizontal organisational structures instead of the traditional top-down hierarchy. According to Schildt (2017), the increasing importance of AI technologies for companies’ operations will shift leadership and power from top management to a wider base of professionals who master programming and data analytics. Utilising AI for this kind of augmenting purpose would mean that an individual employee or leader/manager could accomplish more supported by

technology than the professional or AI software on their own (Jarrahi 2018). Hence, combining trust and technological capacity to augment employees could be the way forward, instead of using AI to monitor and control. However, organisations may find it hard to resist the temptation to use technology to monitor their workforce, especially with the rapid expansion of remote work after COVID-19.

It seems that the most suitable use of technology in the leadership context would be not as the master but in the augmenting, servant sense – to enable and inform employees and leaders. Leadership skills are most needed in novel situations, such as the COVID-19 pandemic – exactly the type of situation in which algorithms struggle. Few would be brave enough to let an AI system lead the way through such uncharted territories. Nonetheless, even in unexpected situations, an AI system could aid human leaders by acting as their artificially authentic peer. What that kind of AI-based management system looks like could be significant.

The vignettes above illustrate how companies may choose from a variety of approaches in utilising AI systems to lead and manage their workforce. As algorithms follow commands with blind obedience, the importance of the organisations' goals, and the rules they follow, are amplified. In recognition of the risks, there have been calls for more accountability of firms, warranties for algorithms or, as argued by Haenlein and Kaplan (2019), a programmer's version of the Hippocratic Oath. What accountability in terms of AI-based leadership would mean, or what effect a programmer's Hippocratic Oath would have on the development of algorithms, is something future research could address.

It may be that only the limitations of technology will determine an AI-based leader's performance in terms of authenticity. Currently, even the most advanced AI systems do not *understand* their decisions or their context, which is crucial in situations involving human interaction. If an employee is underperforming because of personal issues, an AI leader may neglect the context and simply fire the person, while a human leader would ideally opt for a supportive solution. For an AI system, self-awareness (Neider and Schriesheim 2011) is the most incompatible aspect in terms of authenticity. The formula for consciousness may remain a secret for centuries, for humans and AI-based creations.

Implications

The realities of work in modern organisations make it difficult to exhibit authentic leadership. Being authentic as a leader has been associated with different qualities, typically with features resonating with good leadership. Leadership behaviour and authenticity may require another viewpoint. Human shortcomings may be more important for authenticity than previously thought. Instead of portraying a facade of good leadership to appear authentic, a leader needs to be true to what the situation demands and the leadership behaviour their followers need and expect from them. This applies to human and AI-based leaders alike, including the form of appearance (i.e., an avatar, in this case).

Research suggests (e.g., Darling 2015) that if an AI-based leader had an embodied appearance, virtual or robotic, employees are more comfortable interacting with it. Projecting human traits and characteristics on to non-human beings – anthropomorphism – combined with the fact that technology is advancing in leaps and bounds could mean that we may be interacting with AI naturally and with ease sooner than we expect. Even with life-like virtual avatars, it remains challenging for AI-based leaders to fulfil heightened authenticity

expectations, visually or behaviourally. An AI-based leader that looks like a person, but lacks understanding and behavioural authenticity, would be even more frustrating than a non-embodied representation. Exploring the role of expectations in how comfortable we are in human–technology interaction, and how the 'form' of technology shapes those expectations, is, however, intriguing. Therefore, we suggest anthropomorphism as one aspect of future studies on the authenticity of AI-based leadership.

The supposed requirement to be in constant control and take charge requires leaders not to appear weak, even if overwhelmed by a situation. Would this not result in the opposite of authentic leadership: artificial leadership abolishing the 'undesirable' qualities of authenticity and boosting the 'desirable' ones? This would be akin to resorting to artificial sweeteners to avoid the undesired side effects of sugar, which may have unforeseen consequences for the human body in the long run. The same may apply to AI-based leadership in search of an overly authentic experience.

To avoid misguided AI development in authentic leadership, we should probably avoid constructing artificially sweetened 'Zero AI leadership' – similar to Coke Zero – able to simulate authenticity but lacking contextual understanding or common sense. While AI management systems based on machine learning programmes can carry out some management tasks well, we are far from an AI leader being able to lead teams or organisations. For the foreseeable future, we view AI systems as a resource that can augment – but not replace – authentic human leaders.

Analogically to the field of medicine, where AI functions as a diagnostic aid, it could serve leadership like an assistant aiding in diagnosis. As in the medical profession, where the final

treatment call resides with the medical professional, appropriate leadership behaviour requires a human leader with contextual understanding and judgment. This is crucial to ensure that we do not unwittingly end up in the worst-case scenario of an Orwellian dystopia.

Conclusion

This chapter expanded a critical area of research that had previously been highlighted in Horizon Europe Pillar 2 Cluster 4, 'Digital Industry and Space'. While that work did not explicitly reference 'leadership', its key topics pertained to issues raised in this chapter and indicated the wealth of relevant research topics that will impact the nature of leadership moving forward: digital technologies; smart networks and services; high-performance computing; AI-data-robotics. These topics may challenge concepts and applications of leadership as well as the meaning and relevance of authentic leadership. This chapter focused on the extent to which it may be possible to create an AI-based authentic leader, questioning the inherent contradiction between 'artificial' and 'authentic'. It posed the following central research question: Does the concept of authenticity have value in the context of AI? This could also be turned around as follows: Does the application of AI – even just as a powerful resource – challenge the tenets of authentic leadership because the driver of leaders' decisions would effectively be the intelligence and analysis supplied by the AI? This field has enormous potential for interdisciplinary research with studies that integrate case studies, theory development and even AI development. Implicit in this discussion are other potential research topics, such as, for example, Schildt's (2017) assessment of the extent to which knowledge, high-level understanding and competence in the application of AI impacts power relations in an organisation and enhances distributed leadership. If AI plays a 'servant' role, should this dimension of servant leadership be explored?

One potential research avenue that has been enabled by our investigation of the intersection of authenticity, AI and leadership is to assess whether the search for authenticity in AI-based leadership could lead anywhere. The search for the implications of AI for authentic leadership may entrap us in unresolvable existential and conceptual ambiguity, leading us to lose sight of what matters: authentic leadership. As pointed out by Jean-Paul Sartre (1992, p. 4), 'If you seek authenticity for authenticity's sake you are no longer authentic'.

Acknowledgements

Niilo Noponen has received funding from Jenny and Antti Wihuri Foundation (grant number: 00200244).

References

Ajunwa, Ifeoma, Kate Crawford, and Jason Schultz. 2016. "Limitless Worker Surveillance." *California Law Review*, March, 101–42. https://doi.org/10.15779/Z38BR8MF94.

Alkhatib, Ali, and Michael Bernstein. 2019. "Street–Level Algorithms: A Theory at the Gaps between Policy and Decisions." In *Conference on Human Factors in Computing Systems - Proceedings*, 1–13. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3290605.3300760.

Alvesson, Mats, and Katja Einola. 2019. "Warning for Excessive Positivity: Authentic Leadership and Other Traps in Leadership Studies." *Leadership Quarterly* 30 (4): 383–95. https://doi.org/10.1016/j.leaqua.2019.04.001.

Avolio, Bruce J., and William L. Gardner. 2005. "Authentic Leadership Development: Getting to the Root of Positive Forms of Leadership." *Leadership Quarterly* 16 (3): 315–38. https://doi.org/10.1016/j.leaqua.2005.03.001.

Bernstein, Ethan S. 2012. "The Transparency Paradox." *Administrative Science Quarterly* 57

(2): 181–216. https://doi.org/10.1177/0001839212453028.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. “Language Models Are Few-Shot Learners.” *ArXiv Preprint ArXiv:2005.14165*, May. http://arxiv.org/abs/2005.14165.

Buurtzorg 2016 (shorturl.at/BNRXZ, accessed April 21, 2021)

Crossover. 2020. “WorkSmart Productivity Tool.” https://www.crossover.com/pages/worksmart-productivity-tool. Accessed September 9, 2020

Darling, Kate. 2015. “‘Who’s Johnny?’ Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy.” *SSRN Electronic Journal*, April. https://doi.org/10.2139/ssrn.2588669.

Dzieza, Josh. 2020. “Robots Aren’t Taking Our Jobs — They’re Becoming Our Bosses.” *The Verge*, February 27, 2020. https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon.

Einola, Katja, and Mats Alvesson. 2021. “The Perils of Authentic Leadership Theory.” *Leadership*, March. https://doi.org/10.1177/17427150211004059.

Haenlein, Michael, and Andreas Kaplan. 2019. “A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence.” *California Management Review* 61 (4): 5–14. https://doi.org/10.1177/0008125619864925.

Hannah, Sean T., John J. Sumanth, Paul Lester and Fabrice Cavarretta. 2014. "Debunking the false dichotomy of leadership idealism and pragmatism: Critical evaluation and support of newer genre leadership theories." *Journal of Organizational Behavior*, 35 (5): 598–621.

Harms, Peter D., and Guohong Han. 2019. “Algorithmic Leadership: The Future Is Now.” *Journal of Leadership Studies* 12 (4): 74–75. https://doi.org/10.1002/jls.21615.

Harter, Susan. 2002. *Authenticity.* In C.R. Snyder, Shane J. Lopez (eds.) *Handbook of positive psychology.* Oxford: Oxford University Press, 382–394

Heidegger, Martin. 1977. *The Question Concerning Technology and Other Essays* (trans: W. Lovitt). New York: Harper and Row.

Holt, Kristoffer. 2012. "Authentic Journalism? A Critical Discussion about Existential Authenticity in Journalism Ethics." *Journal of Mass Media Ethics: Exploring Questions of Media Morality* 27 (1): 2–14. https://doi.org/10.1080/08900523.2012.636244.

Höddinghaus, Miriam, Dominik Sondern, and Guido Hertel. 2021. "The automation of leadership function: Would people trust decision algorithms?" *Computers in Human Behavior* 116 (March 2021): 1–14. https://doi.org/10.1016/j.chb.2020.106635.

Iszatt-White, Marian, and Steve Kempster. 2019. "Authentic Leadership: Getting Back to the Roots of the 'Root Construct'?" *International Journal of Management Reviews* 21 (3): 356–69. https://doi.org/10.1111/ijmr.12193.

Jarrahi, Mohammad Hossein. 2018. "Artificial Intelligence and the Future of Work: Human-AI Symbiosis in Organizational Decision Making." *Business Horizons* 61 (4): 577–86. https://doi.org/10.1016/j.bushor.2018.03.007.

Jarrahi, Mohammad Hossein, Gemma Newlands, Min Kyung Lee, Eliscia Kinder, and Will Sutherland. 2021. "Algorithmic Management in a Work Context." *Big Data & Society* Forthcoming. https://www.researchgate.net/publication/351462257_Algorithmic_Management_in_a_Work_Context/citations.

Kaine, Sarah, and Emmanuel Josserand. 2019. "The Organisation and Experience of Work in the Gig Economy." *Journal of Industrial Relations* 61 (4): 479–501. https://doi.org/10.1177/0022185619865480.

Kempster, Steve, Marian Iszatt-White, and Matt Brown. 2019. "Authenticity in Leadership:

Reframing Relational Transparency through the Lens of Emotional Labour." *Leadership,* 15 (3): 319–38. https://doi.org/10.1177/1742715017746788.

Lee, Min Kyung, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. "Working with Machines." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 1603-1612*. New York: ACM Press. https://doi.org/10.1145/2702123.2702548.

Lim, Wootaek, Daeyoung Jang, and Taejin Lee. 2017. "Speech Emotion Recognition Using Convolutional and Recurrent Neural Networks." In *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2016*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/APSIPA.2016.7820699.

Ma, Ning F., Chien Wen Yuan, Moojan Ghafurian, and Benjamin V. Hanrahan. 2018. "Using Stakeholder Theory to Examine Drivers' Stake in Uber." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 1-12*. New York: ACM Press. https://doi.org/10.1145/3173574.3173657.

Mayor, Adrienne. 2018. *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691183510/gods-and-robots.

McCarthy, John, Marvin L Minsky, Nathaniel Rochester, and Claude E Shannon. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." Darthmouth College. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf.

McGregor, D., (1960). Theory X and theory Y. *Organization theory*, *358*(374)

Meyer, J. W. (2016). "Foreword". In Shariff, Jay, J. Steven Ott and Yong Suk Jang (eds.) *Classics of Organization Theory*, 8th edition. Boston, MA: Cengage Learning, ix-xii.

Nandram, Sharda, and Nicole Koster. 2014. "Organizational Innovation and Integrated Care:

Lessons from Buurtzorg." *Journal of Integrated Care* 22 (4): 174–84. https://doi.org/10.1108/JICA-06-2014-0024.

Nehamas, Alexander. 1999. *Virtues of Authenticity: Essays on Plato and Socrates.* Princeton, NJ: Princeton University Press.

Neider, Linda L., and Chester A. Schriesheim. 2011. "The Authentic Leadership Inventory (ALI): Development and Empirical Tests." *Leadership Quarterly* 22 (6): 1146–64. https://doi.org/10.1016/j.leaqua.2011.09.008.

Noponen, Niilo. 2019. "Impact of Artificial Intelligence on Management." *Electronic Journal of Business Ethics and Organization Studies*. 24 (2): 43–50. http://ejbo.jyu.fi/.

Parry, Ken, Michael Cohen, and Sukanto Bhattacharya. 2016. "Rise of the Machines: A Critical Consideration of Automated Leadership Decision Making in Organizations." *Group & Organization Management* 41 (5): 571–594.

Rosenblat, Alex. 2018. *Uberland: How Algorithms Are Rewriting the Rules of Work.* Oakland, California: University of California Press. https://www.jstor.org/stable/10.1525/j.ctv5cgbm3.

Rosenblat, Alex, and Luke Stark. 2016. "Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers." *International Journal Of Communication* 10 (27): 3758–3784. https://doi.org/10.2139/ssrn.2686227.

Samuel, Arthur. 1959. "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development* 44 (1.2): 211–29. https://doi.org/10.1147/rd.441.0206.

Sartre, J.P., (1992). Notebooks for an Ethics. University of Chicago Press.

Synthesia 2022 (https://www.synthesia.io/, accessed Sep 12, 2022)

Schildt, Henri. 2017. "Big Data and Organizational Design - the Brave New World of Algorithmic Management and Computer Augmented Transparency." *Innovation* 19 (1):

23–30. https://doi.org/10.1080/14479338.2016.1252043.

Tarnowski, Paweł, Marcin Kołodziej, Andrzej Majkowski, and Remigiusz J. Rak. 2017. "Emotion Recognition Using Facial Expressions." In *Procedia Computer Science*, 108:1175–84. Elsevier B.V. https://doi.org/10.1016/j.procs.2017.05.025.

Tech Times. 2019. "AI Perpetuating Human Bias In The Lending Space." https://www.techtimes.com/articles/240769/20190402/ai-perpetuating-human-bias-in-the-lending-space.htm Accessed May 5, 2021.

Vincit (2022). This is our story (https://www.vincit.com/our-story, accessed September 15, 2022)

Whelchel, Robert J. 1986. "Is Technology Neutral?". *IEEE Technology and Society Magazine*, 5 (4):3–8. https://doi.org/10.1109/MTAS.1986.5010049

Zimmerman, Michael E. 1990. *Heidegger's Confrontation with Modernity: Technology, Politics, and Art*. Indiana University Press.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30 (1): 75–89.