

## 1 Introduction

### Digital society's techno-totalitarian matrix

*Ismael Al-Amoudi and Emmanuel Lazega*

In William Gibson's *Matrix* trilogy, most humans live Hobbesian lives: solitary, poor, nasty, brutish and short. Privately owned companies exert a de facto monopoly on technology and violence through the use of subservient 'salary-men' and through mastery of expensive technologies for spying on and killing, but also for upgrading and downgrading human beings. Throughout a complex plot, humans of various levels of enhancement are manipulated by artificial intelligences (AIs) that seek to bypass the material safeguards and limitations imposed on them by their human creators and owners. These artificial intelligences appear to have developed some form of consciousness, though one that is very, if perhaps not radically, remote from human consciousness. As the story ends, the uploaded mind of a dead protagonist marries his beloved in the Matrix while AIs start colonising a nearby galaxy.

When the *Matrix* trilogy was published in the mid-1980s, it introduced to mass culture a number of post-human tropes that have haunted our collective imaginaries ever since. The most noted is arguably the eponymous Matrix, an information network that prefigures the development of the World Wide Web. But the *Matrix* trilogy also contains several other themes that inspired not only subsequent science fiction writers and cyberpunk fashionistas but also many of the scientists, engineers, entrepreneurs and intellectuals who invented, designed, marketed and commented on the technologies born at the turn of the 21st century.

Indeed, while conscious machines do not exist in 2018, the questions about their possibility in principle and about the process through which they may emerge remain open (Archer, present volume). While mind-uploading is still a fantasy, an increasing amount of human interaction, including the intimate, happens electronically and via online social networks (Donati, present volume). While interactions with bots and robots remain less baffling than in the *Matrix*, their pervasiveness already raises questions about the emergence of new forms of sociality (Maccarini, present volume). While the domination of a small elite that reaps the benefits of technology is not as stark as in the *Matrix*, current automation trends are largely excluded from public debate and left to a few powerful actors, public and private, who seek to influence rather than inform citizens and their representatives (Morgan, present volume). While artificial intelligence (AI) is not capable of consciously manipulating human beings for its own concerns, it has already started to bear on normative decisions in ways that undermine the ethics of human discussion (Al-Amoudi & Latsis, present volume). Finally, while turn-of-the-century soldiers and hit-men do not benefit from the extraordinary healthcare imagined by Gibson, they can already count on AI-based systems of operational support that multiply combat efficiency at the expense of team oppositional solidarity and personal tacit knowledge (Lazega, present volume), thus generating new organisational models.

The aim of our book, however, is neither to marvel at Gibson's prophetic vision nor to describe the gap that still exists between science fiction and science *tout court*. Our purpose is rather to discuss the social significance of phenomena we can already know. We try to understand how post-human technological developments, and especially AI, have started to transform our human agency but also the basic institutions and organisations that hold contemporary societies together: the family (Donati, present volume) and the household (Maccarini, present volume), but also commercial corporations (Morgan, present volume),

health institutions and organisations (Al-Amoudi & Latsis, present volume), and the military writ large (Lazega, present volume).

Our collective book opens with a broad but reflexive literature review by Douglas V. Porpora on AI and human enhancement. The review indicates that, while books on AI started to appear in the 1960s, the topic reached a peak of popularity in the 1980s and, in spite of a slight decline, has remained fairly popular since then. But Porpora's review also provides insight about what the press has to say about AI. To do so, he has examined all articles published on AI by the *International New York Times* (INYT) over a period of 50 days randomly selected in the year 2017. The articles gathered over that period provide a reasonably representative sample of how AI is discussed in connection with four broad themes: social developments; the economy; innovation and capacities; and the arts.

Porpora's review is doubly useful for our project, both because it provides a refresher about the new artefacts, practices and institutions emerging as we write and because it helps appreciate some of the limits of media discourse on AI. For instance, a number of articles express dismay in the aftermath of an AI's victory over the human Go champion. The articles' dismay is based, however, on the widespread assumption that mastery of Go is indicative of human-like 'intelligence'. Yet, the contribution of Archer to the present volume (Archer, present volume; see also Donati, 2019; Morgan, 2019; Porpora, 2019) suggests otherwise: AIs might indeed reach capacities of equal (if not superior) power and worth as human minds; however, what is specific to and valuable about human minds is not so much their computational capacities as much as their endowment with a first-person perspective (Baker, 2000) and their capacity to identify concerns (Archer, 2000) and subsequently reflect on them (Archer 2007, 2010, 2012). But Archer's contribution to the present volume offers more than a mere philosophical critique of misplaced journalistic dismay; it also describes a plausible process through which an AI (as we know them in 2018) might come to acquire, through interaction with a human being, the essential characteristics of human mind: a first-person perspective, concerns and reflexivity.

One of the INYT streams witnessed by Porpora relates to what he calls social developments and reports on the pervasiveness of artificially 'intelligent' machines in daily life. While most articles assume that the threat of 'technological singularity' is still remote and that machines are not even close to over-ruling the world of humans (Kurzweil, 2005), many articles report on the AI-empowerment of familiar objects (e.g. cars, home appliances) and on the appearance of AI-equipped commercial sexbots, that is, machines specifically designed for their owners' sexual enjoyment. The INYT articles do not seem to notice, however, the significance of these technological developments for human sociality. How does living in a world populated with AIs bear on our human capacity to initiate, foster and steer meaningful relations with others?

Yet, the question of human sociality is at the heart of the chapters written by Andrea M. Maccarini and Pierpaolo Donati in the present book. Maccarini (present volume) asks whether interaction with AI-powered machines might encourage people to prefer 'pure relations' with AI machines that are devoid both of bodily imperfections and of the character flaws so common to humanity: impatience, envy, laziness and so forth.

Donati's chapter also addresses the evolution of human sociality, though through slightly different concepts. Donati posits the existence and relative unity of a *digital matrix* consisting in 'the globalised symbolic code from which digital artefacts are created in order to help or substitute human agency by mediating inter-human relations or by making them superfluous' (present volume, p. 105). But the recent emergence of the digital matrix is not, Donati argues, an innocuous addition to the social world. Rather, it deeply transforms and even *hybridises* social relations and the operations of human minds: 'The hybridised family turns out to be a family in which everyone is "alone, together"'. Relationships are privatised

within the already private space of the family, while the identities of the members travel in the public space of the DM [digital matrix]' (Donati, present volume, p. 111).

Another INYT stream of articles discusses AI's economic implications. Most of these express anxiety about job destruction together with an interrogation about whether or not AI systems will replace or complement human labour. But the press articles also take at face value the estimates produced and circulated by governmental agencies and influential consultancies and think tanks. And this is precisely where Morgan's contribution to the present volume starts: since 'little attention was paid in the press to what the economic models were actually claiming or how they were constructed' (Morgan, present volume, p. 94), Morgan proposes to unpack the analyses and assumptions of the UK Made Smarter Review, an influential positional document on the economic consequences of AI technologies. Among other findings, Morgan shows in detail how the report moves from relatively fragile assumptions to seemingly objective figures and from speculation on a fundamentally open future to claims that 'there is no alternative' but to automate quickly enough to safeguard the competitiveness of national firms, thus pitting nations against each other in what could become a race to the bottom.

The third INYT stream identified by Porpora discusses AIs' capacities and limitations. In the face of AI's victory at Go (mentioned previously), human-like capacities of calculation and even of intuition are no longer humanity's preserve. Yet, as several of the shortlisted articles remind us (following Bostrom, 2016), AI programmes are still highly specialized, and those capable of beating a Go champion are incapable of driving a car and vice versa. The implication is not only that AIs are still poor improvisers in unfamiliar contexts but also that they are arguably remote from having their own moral powers or, as Jim Kerstetter has it, 'The better question might be: how do you teach a computer to be offended?' (Kerstetter, 2017, cited in Porpora, present volume).

Taking stock of AI's moral limitations, Al-Amoudi and Latsis (present volume) ask a slightly different, arguably overlooked but equally important, question: How does reliance on AI affect the capacity of human beings to discuss normative decisions? While their discussion is centred on public health, their findings are relevant to a wider array of industries and normative discussions.

In the same vein, Lazega (present volume) tracks how AI increases the capacities of command and control in organisations to unobtrusively shape interactions and parametrise collective agency between humans. A military template for this extension of AI both analyses real-time information from multiple sources and uses digital tools engineered to apply mathematical models of animal swarms for the management of army units operating under high stress in battlegrounds. This involves homogenising mental maps, anticipating response to enemy moves, manipulating emotional reactions, suggesting courses of action, preventing improvised deeds, and defusing oppositional solidarities. Although this capacity deals with soldiers, it could generate new organisational models for non-military organisations, in line with a tradition of military and war technology that have long shaped society at large (Centeno & Enriquez, 2016).

## **Organisational society: smart machines as agents of further bureaucratisation?**

Organisational approaches are useful, and perhaps unavoidable, when reflecting on contemporary challenges to the human condition. Over the last couple of centuries, Weberian sociologists such as Presthus (*Organizational Society*), Jacobi (*The Bureaucratization of the World*), Stone (*Where the Law Ends*) and Coleman (*The Asymmetric Society*) raised concerns

over the growing importance and even colonisation (Deetz, 1992) of most areas of social life by large private organisations. In the words of Charles Perrow:

The appearance of large organizations in the United States makes organizations the key phenomenon of our time, and thus politics, social class, economics, technology, religion, the family, and even social psychology take on the character of dependent variables. Their subject matter is conditioned by the presence of organizations to such a degree that, increasingly, since about 1820 in the United States at least, the study of organizations must precede their own inquiries.

(Perrow, 1991: 725)

To understand contemporary social change in such organisational societies, two ideal types of organised collective action have been identified: bureaucracy and collegiality (Lazega, 2017, forthcoming). These ideal types, each with its specific formal and its informal dimensions, combine social discipline and productive efficiency; they can be observed in real-life companies, associations, cooperatives, public authorities and so forth. The ideal types of bureaucracy and collegiality help us understand the organisational context of work practices, be they routine or innovative. In this dual logics approach, the bureaucratic model is generally employed to organise collective routine work while concentrating power unobtrusively: command and control at the top and depersonalised interactions among subaltern members. The collegial model, on the other hand, is usually observable in situations requiring collective innovative work with unpredictable output. Through collegial organisation, rival peers self-govern by deliberation and agreements or consensus building and by using personalised relationships and relational infrastructures to manage coordination and cooperation dilemmas.

But the ideal types of bureaucracy and collegiality are seldom present in their pure form throughout any single organisation. Rather, real-life workplaces, communities, markets and societies are replete with combinations of collegiality and bureaucracy. Indeed, organisations that can be called 'bureaucratic' (e.g. airlines) are nonetheless managed by a collegial top-team who maintain highly personalised relationships, and conversely, collegial organisations (e.g. private law firms) typically rely on bureaucratically organised support services which interact in largely impersonal ways.

If a lead is taken from the articulation of personalised collegiality and impersonal bureaucracy, the digitalisation of society can be interpreted as both cause and symptom of further and deeper bureaucratisation of society. Does this mean that impersonal interactions, routines, hierarchies and mass production will increasingly characterise our bureaucratized and technocratic contemporary societies? Contributions to this volume address this and underlying issues at varying levels of generality.

Donati argues that human relations are hybridised and even threatened when they are mediated by digital media and smart bots. We are left, however, with the question of how far the depersonalisation of relations can go. Indeed, is a world with human beings but with no personal relations possible in the first place? Or does the digital world necessarily encompass a combination of impersonal transactions and personalised relationships?

Here, a century of organisational sociology and discussion of the bureaucratic model can help us answer. We may draw, in particular, from Jean-Daniel Reynaud's (1989) theory of joint regulation of collective action. From this perspective, there is one dimension of the organisation of collective action that cannot be routinised and that reflects the limits of bureaucracy: the micro- and meso-political negotiation of the 'rules of the game'. This negotiation fleshes out the normative and moral dimension of action, a process of structural and cultural re/production that is never routine and that escapes the capacities of our very best AIs (AI-Amoudi & Latsis, present volume). Apart from the extreme case of totalitarianism (more on this later in this chapter), organisational members do not assume that complete planification and prediction are achievable or even desirable. Continuous coordination of activities is achievable through

common (though necessarily incomplete; see Al-Amoudi, 2010) rules but also through a collective (if contested) project; through (relatively widely) shared cultural schemes of interpretation; and through (reasonably) congruent moral commitments. But the involvement of all actors, even those most subject to bureaucratic control, in negotiation and sense-making does not mean that all are equal in their capacity to defend their regulatory interests. Indeed, the regulatory process produces its share of winners and losers, so much so that Reynaud insightfully reinterprets change and new norms as broken promises. New rules produced by the regulatory process create losers who need to reorganise their practice and joint activities based on the new rules, which leads to the issue of how to handle these losers in bureaucratic contexts and in more collegial ones. For our book's concerns, this means that digitalisation, robots and artificial intelligence are likely to weaken the capacity of most people to defend their regulatory interests but are nonetheless unlikely to eradicate personal relations from the face of society.

Reynaud's reflections on the winners and losers of changing regulation processes help us understand, from a sociology of organisations perspective, both the anxiety expressed by media reports of AI (Porpora, present volume) and the triumphant optimism (Morgan, present volume) exhibited by management consultancies and governmental agencies that implicitly identify with the interests of large organisations and ambitious entrepreneurs.

But it is also because the development of AI is indissociable from its organisational context that Archer's tale ends in melancholic disenchantment: Homer the human and Ali the AI turned into an Artificial Person (an AP?) are both in the end victims of impersonal bureaucracy. Homer is condemned to wander in the limbo of the emeritus professoriate whereas Ali is condemned, without a crime, to lose the personality he developed over all these years and become a sophisticated, though soulless, traffic computer. And this should not surprise us, since organisational bureaucracy was precisely the context in which the lovely synergy of Ali and Homer started in the first place. Bureaucracy giveth and bureaucracy taketh away.

In his discussion of the co-existence and interaction of humans with non-human entities in multiple spheres of social life, Maccarini (present volume) examines the scope of the synergy ventured by Archer. He views AI as a factor of depersonalisation and suggests there is a continuum from mediation to substitution of human social partners. While the threshold between mediation and substitution is clearly theorised, whether and how it is crossed remains an empirical issue. Maccarini's study focuses on 'processes and forms that may properly be called post-human, that is where hybrid forms of social interactions and relations emerge'. But whereas in Archer's account the machine develops personal emergent powers through repeated interaction with a human who remains relatively unchanged, Maccarini's account presents us with the obverse process: the smart machines (as we currently know them) do not evolve, but human personal powers do. And the result is alarming, as it might well involve increased intolerance and misanthropy.

Among the profound effects that AI, robotics and related innovations are set to have on the way we live, work and perhaps even who we are, Morgan looks at some of the ways the future of work is currently being positioned. He begins from the empirical issue of trends in industrial and service robotics and interrogates their significance. He then builds an ontological critique of claims about a fourth industrial revolution by maintaining a sharp distinction between imminent change and immanent potential. By doing so, he shows how positional documents shape our sense of the future, in effect colonising rather than merely predicting our future.

Taking a lead from Morgan, we might be able to refine further the political significance of hybridised (Donati) and pure (Maccarini) relationships. Their chapters in the present volume indicate that we are producing societies where we depend increasingly on impersonal transactions that have nonetheless a highly personalised basis. That is already the case, for example, when dealing with AI bots that can access our data from a variety of platforms. But the combination of impersonal relations and highly personalised transactions is also politically

significant since it casts doubt on one of the founding promises of early cyber-optimists: that with the internet, the last technical obstacles to direct democracy would be removed. On the face of the analyses produced by Archer, Donati, Maccarini, Morgan and Porpora, it seems instead that, while direct democracy remains a distant fantasy, most citizens are gradually losing whatever power they had to use personalised relationships for institutional entrepreneurship. If this tendency is confirmed, then collegiality would be reserved for collegial elites and oligarchies since, as Reynaud argues, joint regulation and politics cannot be routinised and depersonalised. But the major part of the population could still be prevented from building innovative collective action based on personalised relationships. Indeed, emerging tendencies in organisations show that digitalisation can and does undermine the capacity for collective agency found in the myriad collegial pockets that operate below the highest social and organisational strata.

## **Society's organisations: digital corrosion of normativity and solidarity**

During modernity, collegiality in bureaucratised organisations could survive in at least three kinds of 'collegial pockets'. Firstly, in the top-team of the organisation (whether a company, a government, a charity, etc.) must be collegial because work at the top is never routine, always political, and thus requiring joint regulation. Secondly, in professional segments such as legal, educational or R&D departments, as they involve activities requiring tacit knowledge, expertise and creativity. Thirdly, in workers' defense units such as trade unions at the bottom of the organisation, whose collegiality modern capitalist corporations have long tried to weaken or dissolve for fear of oppositional solidarity, whether work-related or ideological. In short, an organisation can never bureaucratise itself entirely, but the top of the organisation can weaken if not undermine collegial pockets among professionals or through union activity of workers on the ground (Lazega, forthcoming). As it happens, we might be currently witnessing the dying of forms of non-managerial collegiality, that is, not instrumentalised by bureaucratic management, at the lower social and organisational levels that are increasingly subjected to bureaucratic regulation and control (Lazega's chapter in Vol. III of *The Future of the Human* book series).

Al-Amoudi and Latsis explore what they call 'the ethics of AI'. To do so, they start from a technical limitation of AI which operates as a 'black box': while we can know its inputs and its outputs, we cannot fully make sense of its internal workings. But this technical limitation generates a broader social and moral limitation, for how will the introduction of AI affect our communities' capacity to discuss, challenge and decide on the norms governing policy, if AIs cannot provide a humanly understandable account of how they perform normatively laden classifications and suggestions?

Rather than address the problem of normative black boxes in general, they chose to focus their discussion on AI developments that are likely to transform/disrupt health policy defined in broad terms. These include activities of national organisations, regulation of mandatory health insurance, prioritisation of certain patient categories, drugs and treatments by supervisory bodies, public health information campaigns, and entwined strategies of for-profit actors such as private insurers or pharmaceutical companies deciding which drugs they will produce and how they will market them. Doing so allows Al-Amoudi and Latsis to identify and discuss novel normative problems arising from AI's bearing on public health policy's practices. Finally, they speculate on how these current developments could lead to a corrosion of normativity through the subordination of human judgement to machine judgement and through

the stigmatisation of people whose profiles do not match the standards set by AI or the interests of those powerful organisations that will own artificially intelligent machines.

But the digital corrosion of normativity is matched by a digital corrosion of oppositional solidarity that Lazega (present volume) can already witness in the army. In order to understand how AI is likely to extend bureaucracy by further digitalising coordination in the workplace, Lazega examines the case of military units in the battlefield and how digitalisation transforms teams on the ground. He starts from soldiers' work and the mechanics of its digitalisation and their meaning in organisational terms. Doing so allows him to discuss the extension of digitalisation to society at large and the consequences in terms of reinventing specific forms of collective responsibility. It appears that, in a creepy way, the process of digitalisation as bureaucratisation of the battlefield neutralises an old problem faced by armies, that is, team-based oppositional solidarities (Shibutani, 1978). This neutralisation is achieved by AI remote-controlled task performance, in which social actors become potentially subjected to new forms of punitive collective responsibility. The specific example used here is military research on designing and using high performance teams based on the 'swarm template'. Swarm fantasies of collective action that can scale up uniformly are developed by mathematicians and engineers combining artificial intelligence and big data, including social network analysis (SNA), reducing relationships between soldiers to impersonal interactions at the "right" physical and social distance. In many ways, if combinations of AI and SNA provide military management with tools that build efficient teams while neutralising oppositional solidarity, they are able to further bureaucratise collegiality as defined here. High-ranking military bureaucrats want to transform previously personalised relationships into interactions that are impersonal and still be able to retransform interactions into relationships when the *esprit de corps* is needed again. In contrast with previous forms of collegiality that relied on personal relationships (bottom-up collegiality and top-down collegiality), this form can be called 'inside-out collegiality' (Lazega, forthcoming) since personalised and private relationships are observed from the outside and instrumentalised, if not transformed into digitalised interactions, deprivatized, neutralised and steered towards alignment. Bureaucratic management has been dreaming of this magic formula, the 'swarm template', for more than a century. The question raised by this combination of AI, SNA and computational social science is whether the meeting of this mathematical template and this managerial dream will have a dystopian effect on society through social digitalisation and whether new forms of collective responsibility, punitive or not, might emerge as a result.

## **The birth of techno-totalitarian societies?**

It is not difficult to see how the digital interruption of normative discussion discussed by Al-Amoudi and Latsis combined with the digital corrosion of oppositional solidarity studied by Lazega can result in massive asymmetries of power between ruler and ruled. Since normative discussion is difficult, the decisions of rulers remain above cultural critique. And since inside-out collegiality inhibits collective action initiated in subordinate groups, the latter are incapable of mobilising effectively to initiate and sustain resistance. From a historical perspective, this combination is unfortunately not new. Indeed, when studying *the origins of totalitarianism*, Hannah Arendt (1951) identified the central generative mechanism behind the rise of totalitarianism as the melting down of traditional solidarities, which leaves a society of atomised individuals incapable of sustaining critical discussion and incapable of initiating collective action without the will of their supreme leader. But while fully blown totalitarianism rests on the physical elimination of vast numbers of individuals for the regime's survival, the earlier phases are both less bloody and less strikingly blatant to internal observers. These early phases

involve nonetheless a number of social processes that are worryingly fuelled by the contemporary digitalisation of social relations.

Firstly, totalitarianism rests on the prior crushing of the individual private sphere which, in a liberal society, legitimately escapes the ruler's gaze and control. The increasing digitalisation of human exchanges is potentially conducive, however, to this phenomenon. In the countries of the Global North, this mechanism is not fully harnessed by states, which still limit their own powers of surveillance over citizens, though with notable exceptions. The first such exception concerns cases of suspected 'terrorism', a signifier with unclear boundaries that justifies nonetheless the suspension of civil rights in the form of indefinite detention, extradition, torture, dispossession and ostracism for the suspected individual but also for her or his acquaintances (Butler, 2004; 2009). It is indeed particularly worrying that states are wilfully ignoring both arguments of principle appealing to human dignity (Porpora et al., 2013) and instrumental arguments of academics claiming that purely military responses to terrorism intensify resentment, and ultimately terrorism itself, more than they contribute to keeping terrorism in check (Chomsky, 2003).

But in organisational society, states are no longer the most powerful actors, and the abuse of individual privacy by for-profit organisations is equally if not more worrying. It is so both because constant monitoring of employees and customers extends to almost all categories of the population (though not its elites, which escape routine monitoring through expensive tactics deployed in the name of trade secrecy), and because such extension is gradually gaining legitimacy (see Al-Amoudi & Latsis, present volume). Employers, for instance, routinely monitor the Facebook accounts of current employees and job candidates in contexts of high unemployment and low welfare-state security nets (see Blanc & Al-Amoudi, 2013, for a discussion of the illiberal consequences of such contexts). As a result, many a person ends up not only self-censoring the content she or he posts but also excluding those inconvenient friends with whom she or he would otherwise remain digitally connected.

Secondly, the narrowing of the private sphere is synergetic, Arendt argues, with the generalisation and normalisation of systems of collective punishment. As she has it:

The consequence of the simple and ingenious mechanism of 'guilt by association' is that as soon as a man is accused, his former friends are transformed immediately into his bitterest enemies; in order to save their own skins, they volunteer information and rush in with denunciations to corroborate the nonexistent evidence against him; this obviously is the only way to prove their own trustworthiness.

(Arendt, 1951: 423)

While the online behaviour of self-censoring employees and job seekers can be interpreted as resulting from guilt by association, another, perhaps more developed, phenomenon can be observed in the financial sphere. Indeed, emergent practices of money-lending based on 'social collateral' (Karlan et al., 2009) are intended to produce individual discipline through fears of collective punishment. Paradoxically inspired by the Grameen approach to microfinance, banks are increasingly using private/intimate data on personalised relationships between neighbours and family to generate ad hoc social units and impose forms of collective responsibility in exchange for lending them money to acquire homes. Combined with other data (financial, professional, legal, etc.), relational data helps evaluate whether or not the lender can rely on these relationships to create enough collective responsibility to repay the loans. In both previous examples, collegiality is observed and manipulated for its own ends by an outside party, a principle that can be extended to all sorts of collegial pockets operated inside out by a ruler with access to the digital information produced by group members.

Thirdly, the rule of law as a source of legitimacy is replaced by the will of the leader. In the case of Bolshevism and Nazism, this was achieved through state propaganda and through the creation of cells that adopted 'so many organisational devices of secret societies without ever



trying to keep their own goal a secret' (Arendt, 1951: 493). The members of such cells were trained to despise existing institutions and to disbelieve the news conveyed by external media. In such a context, respect for the rule of law and for traditional moral norms eroded, and the only source of coordination became the ever-changing will of the movement's supreme leader. It has become an intellectual banality, at least since the democratic failures of the Trump election in the USA and of the Brexit in the UK, that digital social networks generate informational bubbles in which fake news flourishes and resentment grows exponentially.

Less trivial, perhaps, are the authoritarian implications of the introduction of AI examined by Al-Amoudi and Latsis and of swarm-teams studied by Lazega in the present volume. Since the rule of law cannot operate without continuous interpretation by those subject to it (see the discussion of J-D. Reynaud previously), the unfortunate fact that AIs operate as normative black boxes is likely to restrict the capacity of human communities to rely on thoughtfully agreed rules. It is not clear, however, whether the resulting configuration will be one in which authority rests with AI systems' creators and owners or whether the latter will lose control of their creations and will in turn be subjected to anormative AI decisions on topics that should nonetheless require democratic normative discussion. In either case, unless the safeguarding institutions and organisations called for by Al-Amoudi and Latsis at the end of their chapter are implemented, members of digital society are likely to obey unjustified – and therefore authoritarian – decisions rather than the rule of law.

The swarm-teams studied by Lazega in the present volume indicate, however, an even more immediate danger. Oppositional solidarity, be it at the level of teams, organisations or societies, presents a healthy obstacle to autocratic command and control by elites and a condition for the promotion of innovation. And yet, Lazega argues, this mechanism is already disappearing in the army and is bound to disappear in other spheres of society, including educational organisations (teachers being told at the last minute what to teach, to whom and how), in political movements (activists being told what to protest for or against, when and how), in hospitals (doctors and nurses being told who to care for, how and when), and so on.

If these reflections are given credence, then digital society bears the seeds for forms of totalitarianism that could perhaps be less bloody, but also more efficient and thus more stable, than Bolshevism and Nazism. The ontological distinction between imminent events and immanent possibilities (Morgan, present volume) as well as the possibility, in principle, that AIs develop moral consciousness and even forms of thoughtful solidarity with humans (Archer, present volume) provide rays of hope. However, without a public discussion of the social and political implications of digitalisation of societies, organisations and institutions, we are in danger of collectively writing the prequel to Gibson's *Matrix* trilogy.

## References

- Al-Amoudi, I. (2010). Immanent non-algorithmic rules: An ontological study of social rules. *Journal for the Theory of Social Behaviour*, 40(3): 289–313.
- Al-Amoudi, I., & Latsis, J. (present volume). Anormative black boxes: Artificial intelligence and health policy. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Archer, M. S. (2000). *Being Human: The Problem of Agency*. Cambridge: Cambridge University Press.
- Archer, M. S. (2007). *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press.
- Archer, M. S. (2010). *Making Our Way Through the World*. Cambridge: Cambridge University Press.

- Archer, M. S. (2012). *The Reflexive Imperative in Late Modernity*. Cambridge: Cambridge University Press.
- Archer, M. S. (present volume). Considering artificial intelligence personhood. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Arendt, H. (1951). *The Origins of Totalitarianism*. New York, NY: Meridian Books.
- Baker, L. R. (2000). *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.
- Blanc, S. M-A., & Al-Amoudi, I. (2013). Corporate institutions in a weakened welfare state: A Rawlsian perspective. *Business Ethics Quarterly*, 23(4): 497–525.
- Bostrom, N. (2016). *Superintelligence: Plans, Strategies, Dangers*. New York, NY: Oxford University Press.
- Butler, J. (2004). *Precarious Life: The Powers of Mourning and Violence*. London and New York: Verso.
- Butler, J. (2009). *Frames of War: When Is Life Grievable?* London: Verso.
- Centeno, M. A., & Enriquez, E. (2016). *War and Society*. Cambridge: Polity Press.
- Chomsky, N. (2003). Wars of terror. *New Political Science*, 25(1): 113–127.
- Deetz, S. (1992). *Democracy in an Age of Corporate Colonization: Developments in Communication and the Politics of Everyday Life* (SUNY Series in Speech Communication). Albany, NY: State University of New York.
- Donati, P. (2019). Transcending the human: Why, where, and how? In I. Al-Amoudi & J. Morgan (Eds.), *Realist Responses to Post-Human Society: Ex Machina*. London: Routledge.
- Donati, P. (present volume). The digital matrix and the hybridisation of society. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Karlan, D., Mobius, M., Rosenblat, T., & Szeidl, A. (2009). Trust and social collateral. *Quarterly Journal of Economics*, 124: 1307–1361.
- Kersteter, J. (2017, June 16). Daily report: The limits of AI. *The New York Times*. Available online: [www.nytimes.com/2017/06/16/technology/facebook-the-limits-of-ai.html](http://www.nytimes.com/2017/06/16/technology/facebook-the-limits-of-ai.html) (accessed 21 December 2018).
- Kurzweil, R. (2005). *The Singularity is Near: When humans transcend biology*. Viking.
- Lazega, E. (2017). Networks and commons: Bureaucracy, collegiality and organizational morphogenesis in the struggles to shape collective responsibility in new sharing institutions. In M. S. Archer (Ed.), *Morphogenesis and Human Flourishing* (Vol. V, pp. 211-237). Dordrecht: Springer.
- Lazega, E. (forthcoming). Bottom-up collegiality, top-down collegiality, or inside-out collegiality: Research on intermediary-level relational infrastructures as laboratories for social change. In G. Ragozini & M. P. Vitale (Eds.), *Challenges in Social Network Research*. Dordrecht: Springer.
- Lazega, E. (forthcoming). Bureaucracy, collegiality and social change, Cheltenham: Edward Elgar Publishers.
- Lazega, E. (present volume). Swarm-teams with digital exoskeleton: On new military templates for the organizational society. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Maccarini, A. M. (present volume). Post-human sociality: Morphing experience and emergent forms. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.

- Morgan, J. (2019). Yesterday's tomorrow today: Turing, Searle and the contested significance of artificial intelligence. In I. Al-Amoudi & J. Morgan (Eds.), *Realist Responses to Post-Human Society: Ex Machina*. London: Routledge.
- Morgan, J. (present volume). Stupid ways of working smart? Colonising the future through policy advice. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Perrow, C. (1991). A society of organizations. *Theory and Society*, 20(6): 725–762.
- Porpora, D. (2019). Vulcans, Klingons, and humans: What does humanism encompass? In I. Al-Amoudi & J. Morgan (Eds.), *Realist Responses to Post-Human Society: Ex Machina*. London: Routledge.
- Porpora, D. (present volume). What they are saying: About artificial intelligence and human enhancement. In I. Al-Amoudi & E. Lazega (Eds.), *Post-Human Institutions and Organisations: Confronting the Matrix*. London: Routledge.
- Porpora, D. V., Nikolaev, A. G., May, H. M., & Jenkins, A. (2013). *Post-Ethical Society: The Iraq War, Abu Ghraib, and the Moral Failure of the Secular*. Chicago, IL: University of Chicago Press.
- Reynaud, J-D. (1989). *Les règles du jeu: L'action collective et la régulation sociale*. Paris: Armand Colin.
- Shibutani, T. (1978). *The Derelicts of Company K: A Sociological Study of Demoralization*. Berkeley, CA: University of California Press.