1

The complex systems of AI

Recent trajectories of social theory

Anthony Elliott

One of the principal aims of *The Routledge Social Science Handbook of Artificial Intelligence* is to provide students and teachers with a comprehensive and accessible guide to the major topics and trends of research in the social sciences of artificial intelligence (AI), as well as to survey how the digital revolution – from supercomputers and social media to advanced automation and robotics – is transforming society, culture, politics and economy. A unique integration of social science on the one hand and new technologies of artificial intelligence on the other, this handbook offers readers new ways of understanding the rise of AI and its associated global transformations. Another aim of the *Handbook* is to address the very wide array of phenomenon associated with the digital revolution, providing the most up-to-date coverage of developments in AI, machine learning (ML), robotics and supercomputing. Topics addressed where AI currently transforms or, in the future, promises to transform social, economic, cultural and political processes include:

- AI within discrete apps, embedded within operating systems, and operating systems based on AI.
- Single-purpose robots throughout home and work life.
- Bigger, faster, superdata analytics: 'Colossal data' (bigger big data) which will necessarily involve new data curation and analysis approaches that enable more patterns from an ever diversifying range of data.
- Low-power computational hardware, including neuromorphic computers that are more suited to some applications of AI than traditional computers.
- Miniaturized quantum encryption devices, which will underpin the security and trust that will be required before new technologies are widely applied. This particularly applies to applications with high-consequence failure modes (such as implants with direct access to the brain).
- Advances in ML.
- Advances in battery technology: enables stand-alone and mobile intelligence in a wide range of applications.
- Advances in machine cognition systems

- Mainstreaming of human–machine interfaces. This would enable a host of new applications, the easiest to imagine being those using new brain-machine interfaces.
- Massively parallel computational architectures and quantum computing.
- Advances in generalized robotics, such as multipurpose labourer robots.
- Advances in AI, such as the ability to mimic (and improve on) many aspects of human brain function.

The *Handbook* provides representative coverage of the full range of social science engagements within the AI revolution, from employment and jobs to education and new digital skills to automated technologies of military warfare and the future of ethics. A principal aim of the work is to help cross C.P. Snow's 'great divide' – in this instance, that between technical specialists and social scientists on the topic of AI.

A globalizing world of AI

As the great wave of digital technology breaks across the world, artificial intelligence creeps increasingly into the very fabric of our lives. From personal virtual assistants and chatbots to self-driving vehicles and telerobotics, AI is now threaded into large tracts of everyday life. It is reshaping society and the economy. Klaus Schwab, founder of the World Economic Forum, has said that today's AI revolution is 'unlike anything humankind has experienced before'. AI is not so much an advancement of technology but rather the metamorphosis of all technology. This is what makes it so revolutionary. Politics change dramatically as a consequence of AI. Not only must governments confront head-on the fallout from mass replacement of traditional jobs with AI, algorithms and automation, they must ensure that all citizens are adaptable and digitally literate. It will be fundamental to almost all areas of policy development.

Recent technological breakthroughs have resulted in advanced AI transforming manufacturing, the service industry and business platforms, impacting significantly on most jobs including many professions seemingly immune from digital disruption. Research in the US, UK, Japan and Australia, including both academic reviews and government inquiries, estimates that approximately 40% to 50% of existing jobs are at risk from AI technology and automation in the next 15 to 20 years. Other researchers point to a trend of increasing job polarization accompanying automation. At the same time, it has been estimated that AI could contribute approximately $16 trillion to the global economy by 2030.

Given the intricate interconnections between employment and self-identity, it is easy enough to see why more and more people are troubled by AI. Artificial intelligence is, in short, quickly changing the global economy and, fundamentally, everyday life and the self. Smart algorithms run large tracts of enterprise, executing trades, controlling new additive manufacturing, billing clients, automating customer services, navigating aviation flight paths and guiding surgical care. While there is a public fascination with chatbots and self-driving cars, however, very few people understand how AI actually functions and is changing the world in front of their very eyes. Or maybe this is the issue: AI, like electricity, is *invisible*. It is a general-purpose technology that works its magic behind the scenes. The contours and consequences of AI remain elusive to us – we can't see them in action, but we still somehow experience the impact. Like other general-purpose technologies, such as the internal combustion engine, telephony and the silicon chip, AI is becoming ubiquitous. It is everywhere and nowhere at once, both omnipresent and unnoticed.

Whilst there is a lack of agreement among researchers about how to characterize the main defining elements of AI and its related technologies,[1] there is some measure of agreement in the area of public policy and governance. The UK government's 2017 'Industrial Strategy White Paper', for example, defines AI as 'technologies with the ability to perform tasks that would otherwise require human intelligence, such as visual perception, speech recognition, and language translation'.[2] It is perhaps useful to begin with such a definition, one geared to state-promoted AI, if only because such an account is clearly quite narrow, and leaves unaddressed some of the most important deep drivers of AI. It is crucially important, for instance, to underscore the intricate interconnections between AI and ML. A key condition of AI, one not captured by the UK government's white paper, is the capacity to learn from, and adapt to, new information or stimuli. Among the deep drivers of AI are technological advances in the networked communications of self-learning and relative autonomy of intelligent machines. These new systems of self-learning, adaptation and self-governance have helped to reconstitute not only the debate over what AI actually is, but have also impacted the relationship between artificial and organic intelligence.

While AI generates increasing systems of interconnected self-learning, it does not automatically spawn a common set of human reactions or values in terms of those engaging with such technologies. The relation between AI and its technologies, including particularly people's experiences of or views on AI, is a complicated one. As a first approximation we can define AI, and its related offshoot ML, as encompassing any computational system that can sense its relevant context and react intelligently to data. Machines might be said to become 'intelligent', thus warranting the badge 'AI', when certain degrees of self-learning, self-awareness and sentience are realized. Intelligent machines act not only with expertise but also with ongoing degrees of reflexivity. The relation between AI and self-learning is considered to operate at a high level when intelligent machines can cope with the element of surprise. After all, many ML algorithms can easily be duped. Broadly speaking, AI can be said to refer to any computational system which can sense its environment, think, learn and react in response (and cope with surprises) to such data-sensing.[3] AI-related technologies may include both robots and purely digital systems that employ learning methods such as Deep learning, neural networks, pattern recognition (including machine vision and cognition), reinforcement learning, and machine decision-making. Let us take a closer look at some of these approaches and technologies.

Machine learning

Machine learning is one of the most important advancements of contemporary AI technologies, where computers execute tasks through processes of 'learning' or 'information gathering' that draw from (but are not reducible to) human intelligence and human decision making. 'Machine learning', writes Toby Walsh,

> is an important part of computers that think. It tackles the *bottleneck problem*, the problem of pouring into a machine all the knowledge we have developed over thousands of years. Programming all that knowledge ourselves, fact by fact, would be slow and painful. But we don't need to do this, as computers can simply learn it for themselves.[4]

Through analysis of massive volumes of data, ML algorithms can autonomously improve their learning over time. ML relies on algorithms ranging from basic decision trees through to artificial neural networks that classify information by mimicking the structure of the human brain. The rise of neural networks, a kind of ML loosely modelled on the structure of the human brain, consisting of deeply layered processing nodes, has been especially significant in the spread and efficacy of AI. So too, deep leaning – a more recent spin-off of neural networks – which deploys multiple layers of AI to solve complex problems has underpinned much of the explosion of interest from businesses, media, the finance sector and large-scale corporations. The essential scientific aspiration here has focused on replicating general intelligence, which for the most part has been understood largely in terms of reason, cognition and perception, as well as planning, learning and natural language processing.

Natural language processing

Natural language processing (NLP) is a fundamental aspect of AI and encompasses all AI technologies related to the analysis, interpretation and generation (of text- and speech-based) natural language. NLP has prominent applications including machine translation (such as Google Translate), dialogue systems (including Google's Assistant, Apple's Siri and Amazon's Alexa) and automatic question answering (for example, IBM's Project Debater). NLP has matured rapidly over the past 10 to 15 years as a result of the unprecedented amount of language being produced, shared and recorded in electronic and spoken forms.

The social impacts of NLP as conjoined to AI technologies have been massive, and the likely trajectory of development is set to skyrocket. From Amazon's Alexa to Google's Home, people are busy talking to intelligent machines as never before. It is estimated that more than 60% of Internet traffic is now generated by machine-to-machine, and person-to-machine, communication. IT advisory firm Gartner has predicted that by the mid-2020s the average person will be having more conversations with chatbots and robots powered by NLP than with their partner. These claims may seem the stuff of science fiction, but they spell significant change as regards society, culture and politics. I have previously looked at these developments in some detail, focusing on the likely impacts to social interaction and transformations in communication and talk. My argument was that digital devices deploying NLP programs and AI technology are plainly quite divergent from the ordinary conversations of people. Machine talk occurs as part of pre-programmed sequences built up through machine learning. As a result, machine talk – to date at any rate – can usually only respond to conversational contingencies in quite minor ways. Digital devices might be programmed to convey an impression of 'immediate talk' geared to the needs of the user, but the production of machine talk is, in fact, drawn from an enormous database of code, scripted utterances and network conversation. For example, most chatbots and virtual personal assistants consist of programmed 'appropriate replies' to even the most obscure conversations. This is underscored by Brian Christian's argument that machine language is a kind of *conversational puree*, a recorded echo of billions of human conversations. But even this is now under challenge as a result of technological breakthroughs in AI and NLP: for example, Google's Duplex. Chatbots, softbots, and virtual personal assistants have become increasingly integral to our daily lives and our identities, even if we are not always aware of their role. If talking to chatbots and virtual personal assistants becomes the new normal, we should be aware of the ways they could change how we talk to each other and how we relate to ourselves. One thing is certain. AI is having a profound impact on experiences of the self, what identity means, and of how selfhood intersects with others (both human and non-human) in the wider world.

NLP advances and breakthroughs over the past decade have been achieved with specific tasks and datasets, which are driven largely by big data. However, NLP is only ever as good as the dataset underpinning it. If not appropriately trained and ethically assessed, NLP models can accentuate bias in underlying datasets, resulting in systems that work to the advantage of some users over others. Significantly, NLP is currently unable to distinguish between data or language that is irrelevant and socially or culturally damaging. These are matters of significant social and political importance.

Robotics

Robotics has been characterized as the intelligent connection of perception to action in engineered systems. Robotics include not only human-like robots but any kind of technological system that uses sensors such as cameras, thermal imagers or tactile and sound sensors to collect data about the operational environment and construct an automated response-world of actions.

The scaling up of robotics today is hugely significant throughout much of the world. Industrial robots transforming manufacturing – from packaging and testing to assembling minute electronics – is the fastest growing source of robotic technologies. From the early 1960s when one of the first industrial robots was operationalized in a candy factory in Ontario through to the 2010s where new technologies facilitated robots working hand-in-hand with workers, there has been a growing expansion in robotics and the number of published patents on robotics technology. The number of industrial robots in the US jumped from 200 in 1970 to 5,500 in 1981 to 90,000 in 2001.[5] In 2015, the number of industrial robots sold worldwide was nearly 250,000; industrial robotics is an industry which annually enjoys global growth of approximately 10%. Automotive and electronics have been the major industry sectors for robotics use, but many other sectors are increasingly adopting robotics and technological automation. Robotics coupled with converging mobile technologies are especially transforming industry in Asia, which has dominated the ramp-up of robotics use, and with China being the primary contributor. But demand for greater productivity, mass customization, miniaturization and shorter product life cycles has also driven growth for robotics worldwide, especially in Japan, Germany, Korea and the US.

Complexity, complex digital systems and AI

AI does not exist in a vacuum. AI technologies are always intricately interwoven with social systems, as well as other technological systems. Human action (its unfolding and flows), as well as the production and reproduction of social practices, takes place today in the context of complex, powerful technological and social systems that stretch across time and space. The systematic properties of technology and society, specifically their ordering features, give a certain degree of 'solidity' to social practices which are self-organizing, adaptive and evolving. From this angle, technical and social systems are by definition emergent, dynamic and open. Yet such systems are never 'solid' in the sense that they are stable or unchanging. Complex technological and social systems, including the conditions of systems reproduction, are characterized by unpredictability, non-linearity and reversal. The ordering and reordering of systems, structures and networks, as developed in complexity theory, is highly dynamic, processual and unpredictable; the impact of positive and negative feedback loops shifts systems away from states of equilibrium.[6] Drawing from advances in complexity theory, historical sociology and social theory, my contention is that a grounded, theoretically informed account of the digitization of technological and social systems must be based on seven sets of considerations.[7] These considerations are vital to grasping the complex systems of AI. These complex, overlapping connections between technological systems and digital life can be analyzed and critiqued from the sociological considerations I now detail.

First, there is the sheer scale of systems of digitization, of technological automation and of social relations threaded through artificial intelligence – all being key global enablers of the digital data economy. Over 4.5 billion people – more than half the world's population – are online, and digital interactions increasingly impact upon even those who find themselves with limited digital resources.[8] Complex computerized systems of digitization make possible (and are increasingly interwoven with) the production and performance of social life – of business, leisure, consumerism, travel, governance and so on. These systems – of computing databases, codes of software, WiFi, Bluetooth, RFID, GPS and other technologies – make possible our everyday networked interactions, from search engine enquiries to online shopping to social media. These systems facilitate predictable and relatively routine pathways of digitization which underpin smartphone social interactions, online banking, music streaming, status updates, blogs, vlogs and related actions of searching, retrieval and tagging spawned by the Internet. Systems of digitization enable repetition. In the contemporary world of digital life, these systems include social media, CCTV, credit cards, laptops, tablets, wearable computers, uniform resource locators (URLs), smartphones, email, SMS, satellites, computer algorithms, location tagging and so on. The contemporary flourishing of complex, interdependent systems of digitization are the 'flow architectures' that increasingly order and reorder social relations, production, consumption, communications, travel and transport, and surveillance around the world.[9]

Second, AI is not a 'new technology' which simply transcends, or renders redundant, previous technologies. The complex, adaptive digital systems of AI should not be viewed as simply products of the contemporary but, in part, depend upon technological systems which have developed at earlier historical periods. 'Many old technologies', writes John Urry, 'do not simply disappear but survive through path-dependent relationships, combining with the "new" in a reconfigured and unpredicted cluster. An interesting example of this has been the enduring importance of the "technology" of paper even within "high-tech" offices'.[10] Thus, the development and exploitation of digital technologies is interwoven in complex ways with multiple pre-digital technological systems. Another way of putting this is to say that our wireless world is interdependent on a range of wired technologies. Many of the wired technologies – the wires, cables and connections of pre-digital systems – which intersect with digital technologies of WiFi, Bluetooth and RFID date from the 1830s, 1840s and 1850s. There occurred in this historical period an astonishing range of experiments with systems of electrical energy for the purposes of communication. Systems dating from that period based upon the communication potential of electricity include electromagnetic telegraphy (which was trialled in England, Germany and the US in the 1830s), the first viable telegraph line between Washington and Baltimore (constructed by Morse in 1843 with funds from the US Congress), the successful laying of early submarine cables across the English Channel and between England and Ireland in 1851–2 (with a transatlantic cable successfully laid the following decade), and the discovery of the electric voice-operated telephone (demonstrated in 1854 by Antonio Mecucci in New York, although it was some decades later that Alexander Graham Bell conceived the idea for the telephone as a communication system).[11] Subsequent to this period, the twentieth century witnessed a vast array of technological systems emerge and develop. Broadcasting systems – radio from the 1920s, television from the 1940s – were pervasive and hugely consequential for social transformations associated with mass communications. In the 1960s, the launching of the world's first geo-stationary communications satellites spelt the arrival of near-instantaneous communication on a global level. Around this time, other technological systems – from personal computing to mobile telephony – underwent early development, too. The interlinked, tangled dynamics of these 'systems', of which most people are largely unaware as they go about their everyday social activities, is of key importance. Individuals will not know, or entertain awareness of, the conditions, scale or impact of such complex systems since these different technologies fuse and enrich each other.

Third, we need to recognize the global reach of AI as embedded in complex adaptive systems. Whilst the emergence of complex communication networks coincided with the advent of industrialization, it was only in the late twentieth century and early twenty-first century that digital communication technologies and networks were systematically established on a global scale. In this connection, the exceptional significance of various technological transitions that occurred between 1989 and 2007 should be underscored. While digital technologies have progressively developed across time, 1989 was a key time in the constitution of digital life. For this was the year that Tim Berners-Lee invented the World Wide Web through the technological innovations of URL, HTML and HTTP. (The Web did not become readily accessible to people, however, until 1994.) The year 1989 is also significant because Soviet communism collapsed. According to Manuel Castells, this occurred because of Russia's failure to develop new information technologies.[12] Also in this year, global financial markets were increasingly integrated through instantaneous communications and online real-time trading. In addition, mobile telephony was launched, initially through Nokia and Vodafone, through the breakthroughs of the global system for mobile communications (GSM). In 1991, the first GSM phone call was made with a Nokia device through the Finnish network Radilinja.

As the computing technology–inspired 1990s turned into the social media–driven 2000s, the sheer technological brilliance of digitization seemed all the more striking. This next decade ushered in a range of platforms, apps and devices, and along with that the digital transformation of society. In 2001, iTunes and Wikipedia commenced operation. There were also new commercialized forms of social media. LinkedIn was rolled out in 2003, Facebook in 2004, YouTube and Flickr in 2005, and Twitter in 2006. The point, seemingly, was less to apply the digital to everyday life and more to secure one's social niche within the field of the digital. In 2007, smartphones arrived on the market. This was followed by the introduction of tablets in 2010. With the arrival of the 2010s, and such additional platforms as Instagram, Spotify, Google+ and Uber, culture and society was coming to mean status updates, SMS, posts, blogs, tagging, GPS and virtual reality. Digital technologies were transforming social life.

Fourth, there is the sheer ubiquity of AI. I am referring here to various complex, interdependent digital systems which are today everywhere transferring, coding, sorting and resorting digital information (more or less) instantaneously across global networks. With systems of digitization and technological automation, information processing becomes the pervasive architecture of our densely networked environments. As society becomes informationalized as never before, digitization emerges as the operating backcloth against which everything is coded, tagged, scanned and located. Complex automated systems of digital technology emerge as the 'surround' to both everyday life and modern institutions. These technological systems seem to usher in worlds – informational, digital, virtual – that are generalized; that is, these technologies are increasingly diffused throughout contemporary systems of activity and take on the appearance of a functionality which is 'wall-to-wall'. Today's independent informational systems of digitization are, to invoke Adam Greenfield, both 'everywhere and everyware'.[13] From GPS to RFID tagging, and from augmented reality to the Internet of Things, these various interdependent systems are the architectural surround or operational backcloth through which airport doors automatically open, credit card transactions are enabled, SMS is enacted and big data is accessed. As Greenfield contends, this increasingly pervasive digital surround scoops up 'all of the power of a densely networked environment, but refining its perceptible signs until they disappear into the things we do everyday'.[14]

To invoke the possibility of disappearance in this context, as Greenfield does, is to raise the question of the hidden and the invisible as concerns systems of digitalization. Digital life inaugurates a transformation in the nature of invisibility – operationalized through supercomputers, big data and artificial intelligence – and the changing relation between the visible, the hidden and the power. My argument is that the rise of systems of digital technology in the late twentieth and early twenty-first centuries has created a new form of invisibility which is linked to the characteristics of software code, computer algorithms and AI protocols and to its modes of information processing. The invisibility created by digital technologies is that of a protocological infrastructure which orders and reorders the many connectivities, calculations, authorizations, registrations, taggings, uploads, downloads and transmissions infusing everyday life. Codes, algorithms and protocols are the invisible surround which facilitates our communications with others and our sharing with others of personal data through the array of devices and apps and wearable technologies and self-tracking tools that monitor, measure and record people's personal data. The development of WiFi, Bluetooth, RFID and other novel technologies of artificial intelligence has thus created a new form of sociality, based on a distinctive kind of invisibility, which touches on and tracks identities and bodies and constitutes and reorders our social interactions through ubiquitous contactless technologies. But the digital field is, of course, much more extensive in scope, enabling also smart objects (or, anti-wearables) and other digital data-gathering technologies. Many objects and environments have been rendered 'smart' through embedded sensors, interactive visualizations and digital dashboards – again, with an invisible protological infrastructure and the kinds of social relations spawned by it, touching upon the operations of shopping centres, airports, road toll systems, schools and many more.

Fifth, these systems which are ordering and reordering digital life are becoming more complex and increasingly complicated. This growing complexity has powered the rise of ubiquitous computing and AI, and has been underpinned by exponential rates of technological and associated social transformations. 'Moore's Law' has been the guiding maxim of innovation since the mid-1960s, and refers to the so-called doubling of computing power every two years. Computing power is based on the number of transistors in an integrated circuit; and against the backdrop of ever-shrinking computer circuits, engineers have been able to fit exponentially more onto microchips. This has made computers more complex, powerful and cheaper: it is estimated that a smartphone, for example, possesses the computer power previously only available in large mainframe computers. More recently, reports from various technology companies – such as Samsung and Intel – have suggested that beyond 2021 it may not be feasible to shrink transistors any smaller.[15] The limits to technological miniaturization have thus propelled a debate on whether Moore's Law has reached an end point;[16] some analysts argue that quantum computing will provide the new route forward for the continued expansion of computing processing power. And many people believe that ubiquitous computing and AI, when viewed in the context of convergence with nanotechnology, biotechnology and information science, will continue to propel exponential rates of technological complexity, socio-economic innovation and social transformation. Certainly the ubiquity of digital technology, and especially complexity in AI and robotics, involves multimodal informational traffic flows, which in turn substantially depends on technical specialization and complex expert systems.

Sixth, AI technologies go all the way down into the very fabric of lived experience and the textures of human subjectivity, personal life and cultural identities. Complex adaptive digital systems and technological infrastructures are not just 'out there' processes or happenings but are condensed in social relationships and the fabric of peoples' lives. That is to say, complex digital systems generate new forms of social relations as well as reshape processes of self-formation and personal identity. Complex computerized systems, for example, 'bend' social relations towards the short-term, the fragmentary and the episodic – based upon computational interplays of connection and disconnection. 'Life on the screen' (to invoke Sherry Turkle) appears to unfold faster and faster in the early decades of the twenty-first century, as people 'life-splice' the threads of professional, business, family and leisure zones together – using multiple devices across diverse digital platforms. Digital technologies are intricately interwoven with the trend towards DIY, individualized life strategies, where people are busy using devices, apps and bots to schedule and reschedule their everyday lives and experiment with digital life. Systems of digital technology increasingly wrap the self in experiences of 'instantaneous time', and the individualized work of constituting and reinventing digital identities is built out of instantaneous computer clicks of 'search', 'cut-and-paste', 'erase', 'delete' and 'cancel'.

Web-based digital technologies play a constitutive role in social relations today, facilitating digitally downloadable and transferable files containing apps and bots which power the smart-devices that people use 'on the move'. Over 100 billion apps have been downloaded from the Apple App Store alone since 2008,[17] and over 75% of all smartphone users deploy some kind of messaging app – from Facebook Messenger to WeChat to Viber. The instantaneous, just-in-time culture of apps has been a primary conduit through which the great bulk of people in the rich North now communicate, work and socialize. The arrival of the 2020s, however, promises a wholesale shift of social relations into even more accelerated Web-based digital technologies, and specifically the rise of mobile chatbots. This is part of a growing shift to conversational computing, where language is the new user interface which people use for calling upon their digital assistants for booking a hotel room or ordering a pizza. There is already a large online source network of efficient and intelligent bots available for download, and the spread of mobile chatbots looks set to reshape social relationships – both now and into the future.

Seventh, the technological changes stimulated by the advent of complex digital systems involve processes of transformation of surveillance and power quite distinct from anything occurring previously. The expansion of surveillance capabilities is a central medium of the control of social activities – especially the control over the *spacing and timing* of human activities – arising from the deployment of digital technologies to watch, observe, record, track and trace human subjects. From one angle, complex digital systems might be said to have ushered into existence a digital observatory of greatly increased surveillance, somewhat akin to George Orwell's account of Big Brother and Newspeak. Ubiquitous CCTV in public spaces, data-mining software, RFID chips in passports and identity cards, automated software systems governing transport and the speed of vehicles, and the migration of biometric security into various organizational settings: a whole variety of convergent developments has unfolded dramatically extending the scope of digital surveillance. It is evident that the digital monitoring of the citizens' activities and the observing of the online and smartphone interactions of individuals has been undertaken by a growing number of corporations and state agencies. Since former CIA whistle-blower Edward Snowden released documents in 2013 revealing the numerous global surveillance programs run by the National Security Agency with the cooperation of telecommunication giants and various governments, the issue of digital surveillance has moved centre-stage in world politics and, for many critics, has become associated with the production and governance of citizens in the age of neoliberalism. Led by digital technology, the rise of various 'watching technologies' (from CCTV to telerobotics) indicates the arrival of always-on, 24/7 electronic surveillance and a radical extension of the surveillance of subject populations in the political sphere by the modern state.

Critics of digital surveillance tend to be heavily influenced by the late French historian Michel Foucault's notion of panoptic surveillance.[18] Foucault famously identified Jeremy Bentham's panopticon as the prototype of disciplinary power in modernity, and argued that prisons, asylums, schools and factories were designed so that those in positions of power could watch and monitor individuals from a central point of observation. Foucault's panopticon metaphor emphasized the gaze in the sense of surveillance, especially in the form of the continued observation (as in the instance of guards keeping watch upon prisoners or teachers observing a classroom of pupils). These characteristics of disciplinary power have been extended and deepened through digitized surveillance. For example, prisoners can now be kept under 24-hour electronic surveillance. The dispersal of digital technologies of watching are especially consequential for the internalization of surveillance and the more repressive features of disciplinary power. Indeed, some critics understand the digital age as a kind of lifting of the panopticon gaze to the second power, such that digitized surveillance is ever-present and complete.[19]

There can be little doubt that digital forms of surveillance have transformed power relations in contemporary societies, and much more radical developments are likely to result from the next wave of technological innovation. But it is mistaken, I argue, to see digital surveillance as maximizing disciplinary power of the kind described by Foucault. Certainly, some digital systems of surveillance depend upon authoritative forms of monitoring and control, and, in that sense, can be likened to many of the instances of direct supervision discussed by Foucault. But this is not the only aspect of surveillance which comes to the fore in conditions of digital life. Today, surveillance is often indirect and based upon the collection, ordering and control of information. Characteristic of digital interactions such as social media platforms is that there is no centralized location from which individuals are observed; there is instead a distribution of digital interaction across a range of sites and operationalized through a variety of networks. This suggests that the routine use of digital technologies can also be understood in less threatening or menacing ways. Many people now wear self-tracking devices such as Fitbit and Nike's Fuelband, designed to monitor the state of their bodies and provide information on bodily functions such as heart rate, pulse, calories burned and body temperature. New developments in telemedicine make possible the 24/7 monitoring of elderly and vulnerable people; patients who participate in self-care practices through digital monitoring systems are supported by doctors and other healthcare professionals who also access and monitor the health data of patients. Advances in telerobotic technologies within medicine and surgery have been dramatic in recent years, allowing patients in rural and remote areas to access specialist procedures in microsurgery, orthopaedic surgery and minimally invasive surgery in ways not previously possible. Many of the social changes happening to power relations in this technological context cannot be understood as only disciplinary or simply repressive; they also contribute to novel practices of self-care, new forms of selfhood and identity, and the extension of social reflexivity.

Digitized surveillance might perhaps be better characterized as distributed monitoring, a sea of interconnected digital activities ranging from self-tracking to auto-activated information gathering. Central to this idea of distributed informational monitoring, assembled across many platforms and networks, is the notion of 'sousveillance', which refers to people watching each other at a distance through digital technologies.[20] In this digitalization of life, people become part of environments which are sentient and smart, and such digital systems promote increasingly swarming behaviour. Whilst it is acknowledged that professional and personal information is routinely gathered by state agencies through the deployment of digital surveillance technologies, the important point from this perspective is that increasingly indirect forms of surveillance operate 'from below' – as people use digital technologies to click 'like', 'favourite' and 'retweet'. From this angle, people 'watching each other' on social media platforms – Facebook, YouTube, Twitter, Instagram – become caught up in wider processes of surveillance which are at once self-regulating and self-mobilizing.[21]

It follows that another attribute of information monitoring across platforms and networks is that of surveillance at a distance, where data is fluid, decentred, transferred and routinely shared with third parties. As data-mining fast becomes the DNA of the platform economy, one inadvertent, unplanned side effect of the ubiquity of AI has been that complex systems of recording, measuring and assessing the personal information of citizens have become fodder to the business of politics, elections and voting. The 2018 scandal over British political consulting firm, Cambridge Analytica, which harvested data from millions of Facebook profiles to influence voter behaviour in the 2016 US presidential election, is a single example.[22] The data mined by Cambridge Analytica had been contracted through Cambridge University psychologist Aleksandr Kogan; Facebook had previously authorized Kogan to pull data from its online profiles through an app he had developed – *thisismydigitallife* – ostensibly for academic purposes. The app was, essentially, a personality quiz for Facebook users. Before undertaking the quiz, however, users of the app needed to give consent for access to their Facebook profiles as well as the profiles of their Facebook friends. More than 270,000 Facebook users took the quiz, which ultimately resulted in Kogan gaining access to over 87 million Facebook profiles – 30 million of which contained enough information to be matched with other data trails. Cambridge Analytica had invested approximately US$7 million on harvesting this data undertaken by Kogan. Christopher Wylie, a Cambridge Analytica data scientist who became the key whistle-blower on this scandal, commented that this data had been used to construct elaborate psychographic profiles of individual voters. Many commentators argued this it was the data which enabled the Trump campaign to win the electoral college vote while losing the popular vote by three million votes.[23]

The trend towards 'behavioural micro-targeting' of individual behaviour (consumer choices, political affiliations, personal preferences) to 'nudge' or 'steer' election outcomes is part and parcel of the 'dark side' of surveillance in the age of AI. Some critics have argued that there is indeed an emerging system of ubiquitous mass surveillance which is central to the functioning of corporations and governments in contemporary societies. Digital technologies of observation, monitoring, tracking and surveillance of the public and private lives of people function across digital platforms from social networking (Facebook, Snapchat, Instagram) to mobile payment (PayPal, Apple Pay, Google Wallet) to Internet search engines (Google, Yahoo, Bing). Companies use technologies of surveillance to track Web locations, record consumer spending patterns, store emails, manipulate social networking activities and the resulting patterns linked through smart algorithms. 'Facebook', writes Zeynep Tufekci, 'is a giant "surveillance machine"'. The business of surveillance, from the data broker industry to personalized advertising, involves the mining of vast digital data, and the personal information of citizens is routinely bought and sold without the knowledge of the individuals concerned. The result includes major threats to human freedom and privacy, as corporate surveillance over the private and public lives of citizens develops unchecked.

Surveillance is not only a profound structural problem in the digital age; it has been directly marshalled by governments around the world to manipulate and control citizens. Bruce Schneier, in *Data and Goliath*, contends that the ability of governments to peer into our collective personal lives is historically greater than it has ever been:

> Governments around the world are surveilling their citizens, and breaking into computers both domestically and internationally. They want to spy on everyone to find terrorists and criminals, and – depending on the government – political activists, dissidents, environmental activists, consumer advocates, and freethinkers.[24]

Central in many of these surveillance processes is how the state security world deploys data-gathering programs of extraordinary scale, range and depth. For example, the US Prism surveillance operation mines data from Google, Facebook, Verizon, Yahoo and other key Internet companies to track foreign nationals. Similarly, the UK's Government Communications Headquarters (GCHQ) draws data from all Internet and social networking traffic entering Europe to anticipate and prevent cyberattacks, government hacks and terrorist plots. What Louise Amoore terms 'digitized dissection', the disaggregation of a person's data trail into various degrees of security risk, is of key importance to the new surveillance technologies.[25] Such data dissection occurs not only within state borders but also on a global plane. As Schneier concludes, there is today 'a global surveillance network where all countries collude to surveil everyone on the entire planet'.[26] Whilst the advantages to world security of the digital revolution have been considerable, there are clearly many costs stemming from unchecked disciplinary surveillance on citizens. Real dangers include disturbing effects on free speech and freedom of expression, loss of liberty and erosion of democracy.

The Routledge Social Science Handbook of AI

The *Handbook* sets out to provide a reasonably comprehensive account of artificial intelligence – its history, development, deepening, cultures, innovations and technologies – and its relation to the social sciences. The first section of the book examines the ways that social science theory, methods and approaches have responded to, and engaged with, AI. There is an opening detailed discussion of AI from the standpoint of the key concerns of social science, and especially contemporary social theory, by Anthony Elliott. Thomas Birtchnell traces the conditions and consequences of space, place and geography in the discourse of AI. J. Michael Innes and Ben W. Morrison lay out the path of development in the discipline of psychology for encounters with AI, giving insight into the latent conventions and occluded codes that structure psychological life, and indeed the profession of psychology itself, in conditions of advanced AI. Jutta Weber and Bianca Prietl turn to address the discourse of technoscience – the magical terrain of what might be termed 'the technological fix' – in the wider frame of machine learning and predictive analytics. Ross Boyd addresses one of the most central debates in the AI and society canon – namely, how automated intelligent machines are transforming the world of work. Michaela Pfadenhauer and Tobias Lehmann turn to consider the effects of AI, highlighting the complex ways in which personhood and passion are reshaped in the aftermath of the AI revolution. Joffrey Becker discusses AI in the field of anthropology. The sphere of ethics in the wake of AI is analyzed with great insight and sophistication by Vincent Mueller. Finally, Naoko Abe considers developments in design thinking, with a fascinating discussion of human–machine interfaces and the reworking of design and the aesthetic domain in the contemporary era.

The second section of the book shifts focus to address diverse fields of artificial intelligence in social science research. Roman Batko outlines the contributions and consequences of AI to the fields of business and enterprise, focusing on transformations of economy and organization as a result of the digital revolution. Transformations of mobility are centre stage in Sven Kesselring and Carolin Schoenewolf's encounter with AI. Maja De Neergaard and Malene Freudendal-Pedersen look at the interconnections between digitalization and smart cities. Nicolas Petit and Jerome De Cooman turn to consider new developments in law and the legal profession, considering the massive impacts of automation in the legal domain. Matteo E. Bonfanti, Myriam Dunn Cavelty and Andreas Wenger address the vital area of cybersecurity. Frank Sauer writes about some of the most perplexing issues to do with developments in lethal autonomous weapons systems, outlining the ever-increasing connections between AI and the techno-militarization of war. Massimo Durante details his impressive theory of computational power in the age of AI, focusing on how contemporary representations of the world are increasingly recast in the image of machine intelligence. The interconnections between AI, robotics and aged care is the central theme of Eric L. Hsu's discussion of technogenerians. Jo Bates looks at the world of big data, situating the datafication of the world in the context of critical data studies. Entertainment is also reconfigured powerfully as a result of the intrusion of AI into the culture industries, as Sam Han discusses in some detail. Norina Gasteiger and Elizabeth

Broadbent look at the complex relations between AI, healthcare and medicine, with a strong focus on the radical transformations that new digital technologies offer to patient care and well-being. In the final chapter, Louis Everuss provides a superb overview of how AI reshapes migration (both freely chosen and enforced), giving rise to a world of smart borders.

Notes

1 In a review of the first 50 years of AI research, for example, Hamid Ekbia identified eight major science and engineering approaches to AI, each centred on very different characterisations of what 'intelligence' is. See Hamid Ekbia, "Fifty Years of Research in Artificial Intelligence", *Annual Review of Information Science and Technology*, Vol. 44, Issue 1, 2010, pp. 201–247; See also Anthony Elliott, *Making Sense of AI: Our Algorithmic World* (Cambridge: Polity Press, 2021).

2 Department for Business, "Energy and Industrial Strategy, Industrial Strategy: Building a Britain Fit for the Future", November 2017, p. 37, www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf.

3 PwC 2017. Sizing the prize: PwC's Global Artificial Intelligence Study, www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

4 Toby Walsh, *2062: The World that AI Made* (Melbourne: La Trobe University Press, 2018), p. 15.

5 R. Ayres and S. Miller, *The Impacts of Industrial Robots* (Pittsburgh: Carnegie Mellon Robotics Institute, 1981), p. 3; V. Sujan and M. Meggiolaro, *Mobile Robots: New Research* (New York: Nova Science Publishers, 2005), p. 42.

6 See John Urry, *Global Complexity* (Cambridge: Polity Press, 2003).

7 These seven considerations are set out in more detail in my earlier work on AI, and I draw directly from this research in what follows here. See Anthony Elliott, *The Culture of AI: Everyday Life and the Digital Revolution* (London and New York: Routledge, 2019), especially Chapter 1.

8 See "Digital Skills Crisis", House of Commons Science and Technology Committee, UK Parliament, Second Report of Session 2016–17, https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/270/270.pdf.

9 Karin Knorr Cetina, "From Pipes to Scopes: The Flow Architecture of Financial Markets", *Distinktion*, Vol. 7, 2003, pp. 7–23.

10 John Urry, *What is the Future?* (Cambridge: Polity Press, 2016).

11 John B. Thompson, *The Media and Modernity: A Social Theory of the Media* (Stanford: Stanford University Press, 1995), p. 153.

12 Manuel Castells, *The Collapse of Soviet Communism: A View from the Information Society* (Los Angeles: Figueroa Press, 2003).

13 Adam Greenfield, *Everyware: The Dawning Age of Ubiquitous Computing* (Berkeley: New Riders, 2006).

14 Ibid., p. 26.

15 Semiconductor Transistor Association, "International Technology Roadmap for Semiconductors", 2015, www.semiconductors.org/main/2015_international_technology_roadmap_for_semiconductors_itrs/ (Viewed 31 August 2016).

16 There are numerous works which express a sceptical view about the quickening pace of technological advances. Bob Seidensticker's work, *FutureHype*, offers one such sceptical account, but there are of course others. Bob Seidensticker, *FutureHype: the Myths of Technological Change* (San Francisco: Berrett-Koehler Publishers, 2006).

17 www.theverge.com/2015/6/8/8739611/apple-wwdc-2015-stats-update.

18 Numerous authors writing on the theme of surveillance in society, such as Christian Fuchs, have been highly influenced by the work of Michel Foucault. However, it should be noted that there have been numerous calls to recognize the limits of Foucault's theories, especially on the theme of panopticism. See for example Kevin Haggerty, "Tear Down the Walls: On Demolishing the Panopticon", in *Theorising Surveillance: The Panopticon and Beyond*, edited by D. Lyon (Uffculme, Devon: Willan Publishing, 2006), pp. 23–45.

19 See David Lyon, *Surveillance Studies* (Cambridge: Polity Press, 2007).

20 See Rob Kitchin, *The Data Revolution* (New York: SAGE Publications, 2014).

21 For further discussion see Christian Fuchs, "New Media, Web 2.0 and Surveillance", *Sociology Compass*, Vol. 5, Issue 2, 2011, pp. 134–147; See also Samantha Adams, "Post-Panoptic Surveillance Through Healthcare Rating Sites", *Information, Communication and Society*, Vol. 16, Issue 2, 2013.

22 Cambridge Analytica had been established in 2013 as a subsidiary of Strategic Communications Laboratories, an entity which was partly owned by Robert Mercer – an American hedge-fund manager who has strongly backed various conservative political causes. Steve Bannon, then publisher of the alt-right Breitbart News and subsequent advisor to Donald Trump, was vice president of Cambridge Analytica. See Matthew Rosenberg, Nicholas Confessore, and Carole Cardwalladr, "How Trump Consultants Exploited the Facebook Data of Millions", *New York Times*, 17 March 2018, www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html.

23 In a UK TV investigative report, Cambridge Analytica's then CEO, Alexander Nix – who was subsequently suspended by the company – boasted to an undercover reporter about the 2016 Trump campaign: "We did all the research, all the data, all the analytics, all the targeting, we ran all the digital campaign, the television campaign, and our data informed all the strategy. See ABC News (2018) Cambridge Analytica bosses claimed they invented 'Crooked Hillary' campaign, won Donald Trump the presidency, 21 Mar 2018, www.abc.net.au/news/2018-03-21/cambridge-analytica-claimed-it-secured-donald-trump-presidentia/9570690.

24 Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (New York: Norton, 2015), p. 7.

25 Louise Amoore, "Algorithmic War: Everyday Geographies of the War on Terror", *Antipode*, 41, 2009, pp. 49–69.

26 Schneier, *Data and Goliath*, p. 91.