Volume 8 | Spring 2020

DEFENCE STRATEGIC COMMUNICATIONS

The official journal of the
NATO Strategic Communications Centre of Excellence

ISSN: 2500-9486
DOI: 10.30966/2018.RIGA.8

AI ETHICS: A STRATEGIC COMMUNICATIONS CHALLENGE

A Review Essay by Alex Lawrence-Archer

AI Narratives: A History of Imaginative Thinking about Intelligent Machines
Stephen Cave, Kanta Dihal, and Sarah Dillon (eds). Oxford University Press, 2020.

Rage Inside the Machine: The Prejudice of Algorithms, and How to Stop the Internet Making Bigots of Us All
Robert Elliot Smith. Bloomsbury Business, 2019.

Keywords—*strategic communication, strategic communications, AI, artificial intelligence, applied ethics, data science, machine learning*

About the Author

Alex Lawrence-Archer is Chief Operating Officer of the Centre for Data Ethics and Innovation, an independent advisory body set up by the UK Government to advise on the ethics of data and artificial intelligence.

DOI 10.30966/2018.RIGA.8.6.

Artificial intelligence (AI) is increasingly seen as one of the defining technologies, if not *the* defining technology, of our era. In the last decade, it has been under an increasingly intense policy spotlight. Around the world, governments are engaged in a concerted effort to promote, develop, and encourage the use of AI.

No developed nation's industrial policy is complete without an AI strategy. The UK's AI Sector Deal aims to 'put the country at the forefront of the AI revolution'.[1] EU Commission President Ursula von der Leyen made AI one of her top legislative priorities. China's *Next Generation for Artificial Intelligence Plan* aims for the country to become the 'primary' centre for AI innovation by 2030.[2] The OECD AI Policy Observatory lists no fewer than eleven documents within the category 'National Strategy' for the United States.[3] These documents reveal not only a desire to promote the development of AI, but a drive to achieve 'leadership' in AI and to promote innovation 'at home'.

The increased focus on AI among governments has been matched in civil society, academia, and the media, with the establishment of new research institutes and policy programmes, rapid growth in AI course enrolment, and a slew of popular-science books on the topic.

Two recent books—*AI Narratives* edited by Stephen Cave, Kanta Dihal, and Sarah Dillon, and *Rage Inside the Machine* by Robert Elliot Smith—present two quite different conceptions of the term 'AI', and demonstrate that it is being used to cover an extraordinarily wide range of technologies and ideas. Taken together, *AI Narratives* and *Rage* illustrate both why communicating about AI is crucial to achieving strategic AI goals and why doing so is uniquely challenging.

Two letters, many meanings

AI is a term used relatively loosely in policymaking, in the media, and even in academia. It's a conveniently broad term that indicates the general space we are now working in, but its use can mask important nuances.

Historically, AI has often been thought of as the artificial replication of *human* intelligence, as machines capable of reasoning in the same way that we do. Machines, which—even if they are 'enslaved' in some way—are autonomous,

1 AI Sector Deal, UK Government, May 2019.
2 *China Science and Technology Newsletter*, PR China Ministry of Science and Technology, September 2017
3 Policy initiatives for United States, National strategies, agendas and plans, OECD.AI, 2020.

with a will and mind of their own. Part I of *AI Narratives* details the pre-20th century history of thought about this type of AI, from handmaids made of gold appearing in Homer's *Iliad,* to a metal statue capable of reason designed (legend has it) by medieval philosopher Albertus Magnus, and fraudulent speaking dolls that toured 18th century Britain.

AI Narratives has a strong bent towards AI that replicates human-like reasoning and independent will. Part II explores modern AI narratives through several different lenses, including the enslavement of artificial will [Chapter 8], the mindedness of AI [Chapter 10], AI as the 'child' of humanity [Chapter 11], and the possibility of uploading a human mind into a machine [Chapter 13].

AI Narratives therefore focuses on questions and controversies directly related to the idea of human-like AI. These are questions and ideas familiar to us from popular culture—testament to the traditional dominance of human-like AI in our shared imagination: *Is it conscious? Should it have human rights? Will it stop obeying us, outmatch our capabilities, and take over?* Fascinating though these questions are, they have little relevance to the type of AI that has been the focus of advances in capability, deployment, and social change over the past decade.

As forms of AI have moved—through advances in computing—from the realm of fantasy into our everyday reality, so AI as a term has expanded into more of an umbrella term. It encompasses both 'broad AI'—the capacity to understand or learn any intellectual task that a human being can, and 'narrow AI'—non-autonomous AI capable of performing limited, specific tasks very well.

Rage provides a useful counterpoint to *AI Narratives* in that it focuses heavily on narrow AI—the type of AI in use throughout our economy and society today. Specifically, *Rage* charts AI advances that have been made possible by the increasing sophistication and deployment of data science (the extraction of knowledge and insights from structured and unstructured data) and machine learning (a branch of data science in which computer programmes are trained to optimise for a given variable—for example the number of web users clicking on a link in an advert—and make predictions based on large amounts of past data).

That it is a subset of AI and is not meaningfully autonomous does not make narrow AI insignificant. While it is true that we have long used algorithms and data to aid decision-making, recent years have seen a step-change in the amount of data available, the ease (and cheapness) with which it can be collected, stored,

and processed, and the complexity—and sometimes opacity—of the algorithms developed in the process.

As *Rage* outlines, machine learning algorithms and their predictions, recommendations, and decisions have come to be integral to everything from searching the internet to making decisions on paroling prisoners. The same fundamental technological principles underpin image recognition, which will be integral to autonomous vehicles, and natural language processing and voice recognition, which are needed to run the personal assistants on our phones.

Policy initiatives, national strategies, and research and start-up ecosystems, not to mention commercial applications that have taken on strategic significance for developed nations over the past decade, are overwhelmingly concerned with narrow AI. However, this bears little resemblance to the images that historically have dominated the popular imagination—the ideas explored in *AI Narratives*.

This essay goes on to outline why the growing use of narrow AI has come to be seen as an ethical issue, and how this makes strategic communications an important tool in the implementation of national AI strategies.

What do we mean by 'ethics'?

Ethics is a branch of philosophy that deals with questions of right and wrong, of what is morally good or desirable for individuals and society. Ethical principles can tend towards the abstract: accountability, fairness, economic welfare, and human flourishing, for example. Such principles are inherently contestable; the 'right' amount of transparency in a given situation is not an objective matter. And they are often in tension with each other; a solution that uses public money most efficiently may not be the most equitable or transparent.

New technologies—especially those that have a significant impact on society—often raise ethical questions. They change how our world is structured; they create winners and losers. Ethics do not provide an easy formula for answering the questions raised. But because they go to the heart of how society 'should' work, ethics are at the centre of debates about how society responds to major technological shifts.

Why is AI an ethical issue?

Where there are AI strategies, there is AI ethics. The UK's AI Sector Deal refers to the creation of the Centre for Data Ethics and Innovation, established to help ensure safe and ethical innovation in data-driven technology. As part of its AI strategy development, the European Commission has promulgated Ethics Guidelines for Trustworthy AI.[4] Even the US Department of Defence has adopted AI ethics principles.[5]

Ethics are central to national efforts to promote AI for three principal reasons:

1. There is an ethical imperative to seize—and to equitably distribute—the significant benefits that AI technology offers. Recent years have seen major advances in the use of AI to diagnose certain conditions—especially those relying on the analysis of medical scans—opportunities that are vital to seize. Huge economic gains have been made using AI to better match consumer demand with supply through online shopping platforms, but questions about the distributional impact of this change abound.

2. The growing use of AI creates diverse ethical dilemmas in relation to privacy, corporate and state power, transparency, bias, and autonomy to name a few. Online platforms are incentivised to hoover up increasing amounts of our data, often in ways that are not particularly transparent. Algorithms used in financial services and recruitment have been shown to be consistently biased against women due to their reliance on historical data.

3. The growing use of AI is likely to have impacts on democracy and sovereignty. AI systems now mediate the flow of the information that makes up our public discourse, potentially impacting the outcome of elections. And AI promises (or threatens) to open up new technological avenues in warfare and security.

4 *Ethics Guidelines for Trustworthy AI*, Independent High-Level Expert Group on Artificial Intelligence/European Commission, April 2019.
5 C. Todd Lopez, 'DOD Adopts 5 Principles of Artificial Intelligence Ethics', US Department of Defense Blog, February 2020.

DOI 10.30966/2018.RIGA.8.6.

Opportunities

It is relatively uncontroversial to argue that the widespread adoption and implementation of AI technology that already exists promises significant benefits. One recent estimate puts the incremental contribution AI could make to the global economy by 2030 at 16%, or $13 trillion through new products and services, productivity gains, and more efficient matching of supply and demand.[6]

AI also promises to help improve public administration, enabling the state to deliver more and better services more cheaply through improved anomaly detection, demand prediction, and tailoring of services, and more consistent and scalable decision-making in general. Building on behavioural insights and commercial targeting techniques, AI can even be used to influence citizens' behaviour in ways that promote social objectives.[7] Cases exist in almost every sector, from health diagnosis and revenue collection to defence and security.[8]

Perhaps most promising of all, AI has the potential to help us solve some of society's most intractable problems. At the time of writing, AI is being used to respond to the Covid-19 pandemic through drug discovery, outbreak detection, diagnosis, and demand prediction within the health sector.[9] Most controversially, contact tracing via relatively simple AI combined with high rates of smartphone penetration may form part of the approach to scaling back strict social distancing measures in many countries.[10] In the longer term, AI may play a key role in helping us reduce climate change through better climate modelling and optimising energy efficiency at scale.

Ethical dilemmas

Machine-learning-based AI that promises these benefits is reliant on large amounts of data, often opaque, and weighted towards optimising based on past outcomes. These features, among others, have meant its implementation at scale and in increasingly sensitive contexts, which has thrown up significant ethical dilemmas.

6 'Notes from the AI Frontier: Modelling the Impact of AI on the World Economy', McKinsey Global Institute, 4 September 2018.
7 Sofia Ranchordás, 'Nudging Citizens Through Technology in Smart Cities', *International Review of Law, Computers & Technology*, Volume 33 (2019).
8 Alexander Babuta, 'Artificial Intelligence and the Integrated Review: The Need for Strategic Prioritisation', Royal United Services Institute, April 2020.
9 Bernard Marr, 'Coronavirus: How Artificial Intelligence, Data Science And Technology Is Used To Fight The Pandemic', *Forbes*, 13 March 2020.
10 Leo Kelion,'Coronavirus: NHS Contact Tracing App to Target 80% of Smartphone Users', *BBC News*, 16 April 2020.

DOI 10.30966/2018.RIGA.8.6.

These are not the ethical dilemmas of consciousness, independence, and agency that are the main focus of *AI Narratives*. For ethical challenges created by machine learning, *Rage* is a more reliable guide, detailing variously dehumanisation of workers [Chapter 6], bias against marginalised groups [Chapter 7], and increased polarisation in online discourse [Chapter 10].

Rage is just one of a slew of recent books to have explored these and other issues. There is a growing awareness of the characteristics of this technology that seriously threaten to turn people against it, if it is not well-managed:

- The reliance of AI on data and the increasing ease with which it can be collected and processed, creates strong incentives for organisations to gather data on individuals, which creates risks to individual and collective privacy. This dynamic was particularly evident in the recent C*ambridge Analytica* scandal.

- AI is significantly shifting the balance of power between individuals on the one hand, and institutions on the other. From optimising online experiences for the consumer to maximise sales, to subjecting workers to algorithmic management, to tracking citizens using facial recognition technology, companies and states are increasingly able to predict and influence our behaviour in significant ways.

- The fact that machine-learning-derived algorithms are often opaque in their operation—sometimes referred to as 'black boxes'—creates problems in contexts where understanding why a decision has been made is vital for procedural fairness, such as in criminal justice and welfare systems.

- Reliance on data about past decisions and circumstances makes AI prone to replicating the historical biases of human decision-makers. In recruitment, for example, algorithms have been shown to exhibit bias against female candidates, reflecting historical patterns of human hiring managers' decisions. In criminal justice, AI systems have been shown consistently and unfairly to rate black defendants at greater risk of recidivism than their white counterparts in the US.

DOI 10.30966/2018.RIGA.8.6.

Sovereignty

States face an imperative not only to seize the opportunities that AI represents and manage risks to their citizens, but to develop a domestic ability to understand, regulate, and control the technology. Failure to do so means states risk being unable to meaningfully influence technology that increasingly mediates everything, including communications, finance, industrial processes, and warfare. The influence of social media algorithms on how information circulates during election periods is a case in point.

AI ethics initiatives—inside and outside government—have proliferated globally in an attempt to grapple with these issues. It is a broad and rapidly developing area of policy characterised by a sense of urgency as each state races to define its own distinctive approach to the technology, hoping to avoid becoming an AI 'rule- and standards-taker'. In the UK and EU in particular, trust in AI has become one of the core components of that approach.

'Trust' in AI?

The EU White Paper on AI is subtitled 'A European Approach to Excellence and Trust'. The Hall and Pesenti Review, which underpins the UK Government's current AI strategy, cautions that 'building public confidence and trust will be vital to successful development of UK AI'.[11]

Trust in AI's development and use is both an indicator that the ethical challenges of the technology are being well managed, and a key to achieving the scale of adoption necessary to realise its benefits and share them equitably. If using AI fails to win and maintain public trust, we risk a reluctance to adopt the technology, and to use and share data with AI-powered services. That will lead to lower levels of training, and to major challenges in realising benefits in the public sector. There scrutiny is high, officials are risk-averse, and broad social licence to innovate is needed. In the medium-term, low levels of trust incline to a build-up of public pressure, which could result in poorly thought-through regulation.

Creating an environment in which AI is both trustworthy and trusted by the public is vital to AI strategies in the UK, the EU, and further afield. The data-

11 Wendy Hall and Jérôme Pesenti, *Growing the Artificial Intelligence Industry in the UK*, October 2017.

driven Covid-19 contact tracing apps that look set to form part of a number of governments' lockdown exit strategies are thought to need between 60 and 80% penetration among the population to be effective.[12] Short of compulsion, public trust in this technology will be non-negotiable if such levels are to be achieved. Trust in AI is a strategic issue.

Information on the extent to which AI is trusted as a technology is patchy. ODI/YouGov polling in 2019 showed that only 5% in the UK trust social media companies to use their data ethically, rising to only 30% for central government.[13] IPSOS global polling in 2019 found that twice as many people felt that both government and business use of AI should be more tightly controlled, and that 41% of people are generally uneasy about AI.[14]

Even if long-term and consistent measurement of trust in AI and how it is governed is limited, there is good reason to believe that it is not trending in the direction necessary for the long-term success of national AI strategies.

Government action to guide the development and use of AI with a view to making it more *trustworthy* has significantly increased in recent years. The establishment of the UK Centre for Data Ethics and Innovation and the EU Guidelines for Trustworthy AI are just two examples. Similar initiatives have been set up in Singapore,[15] Germany,[16] and Canada,[17] to name a few. For now, this work is relatively behind-the-scenes. Public narratives about AI and how they change in response to initiatives like these will determine whether AI becomes not only more trustworthy, but more *trusted* by the public.

Narratives, trust, and public concerns

A number of historically controversial technologies are useful in demonstrating the role that narratives play in determining levels of trust, adoption, and development.

The UK and EU experience of genetically modified (GM) food technology is frequently cited as a salutary lesson for those seeking to promote the adoption of

12 Chris Stokel-Walker, 'Can Mobile Contact-tracing Apps Help Lift Lockdown?', *BBC News*, 16 April 2020.
13 'Nearly 9 in 10 People Think it's Important that Organisations Use Personal Data Ethically', *The Open Data Institute*, 12 November 2019.
14 'New Global Poll: Widespread Concern about Artificial Intelligence', IPSOS MORI, 1 July 2019.
15 'Artificial Intelligence', Singapore Media Development Authority, 2019.
16 *Report of the German Data Ethics Commission*, Bundesministerium der Justiz und für Verbraucherschutz, 2019.
17 *Responsible Use of Artificial Intelligence (AI)*, Government of Canada, 2019.

DOI 10.30966/2018.RIGA.8.6.

AI in the UK. This new technology promised enormous potential productivity gains, particularly in developing countries. In 2004, Burke detailed how it came to be deeply mistrusted in the UK and the EU, with public narratives focusing on the risks it represented, on potential runaway impacts, and on the lack of consensus among scientists about the technology.[18]

In their earlier work, *Portrayals and Perceptions of AI and Why They Matter,* the editors of *AI Narratives* explain how GM food...

...became a lightning rod for these broader societal concerns. In such narratives, multinational corporations played a key role, often contributing to scepticism about who would benefit from the widespread adoption of GM.[19]

Despite the weight of evidence showing that the technology is safe, by 2014 YouGov polling showed twice as many (40% to 22%) in the UK felt the government should not promote the technology. Although far from universally popular, GM technologies enjoy a greater level of trust in the US. Pew Research in 2016 showed 61% of US adults feel GM food is either better, or no better or worse, than non-GM for one's health. Doubtless, factors other than narratives and trust have had an impact, but since the 1990s both the US and the EU have taken starkly different courses. The US is a world leader in GM crop cultivation; almost no GM crops are commercially grown in the EU.

Advances in human fertilisation and embryology technology provide a complementary example. This ethically controversial and emotive technology promised enormous benefits to individual families and to the understanding and treating of diseases. The mid-1980s saw a concerted, government-led initiative to debate the ethical issues of the technology,[20] culminating in the Warnock Report of 1984, which underpinned the development of a robust regulatory system for the technology in the UK.[21] Public opinion of embryonic stem-cell research, an especially controversial branch of this technology, has been net-positive and steadily increasing since 2003 in the UK.[22]

18 Derek Burke, 'GM food and crops: what went wrong in the UK?', *European Molecular Biology Organization,* EMBO Rep. 2004 May; 5(5): 432–36.
19 *Portrayals and Perceptions of AI and Why They Matter,* The Royal Society, November 2019.
20 Jo Thomas, 'British Debate Embryo Research', *The New York Times*, 16 October 1984.
21 'The Warnock Report on Human Fertilization and Embryology', *Journal of Advanced Nursing*, Issue 10, July 1985.
22 'Stem Cell Research', Gallup, 2019.

These analogies are limited. The technologies concerned are very different. AI is a general purpose technology that anyone can make use of without a lab or (for now) a licence. It is a technology that is already in widespread use; meaning, we face the challenge of 'catching up' to existing narratives. But these examples do, I argue, indicate the impact that popular understanding about a technology's trustworthiness can—over time—have on the extent to which it is developed and adopted.

AI narratives

There is limited documentation of the public debate on and narratives about AI and how it is governed. *AI Narratives* shows how the most dominant narratives in academic and policy literature and in general fiction are those concerned with human-replicating AI rather than with machine learning as we see it today (the *Black Mirror* TV series is a notable exception). In the press, discussions of the potential for automation to lead to large-scale unemployment appears to have achieved significant cut-through.[23]

Rage is part of a trend in popular science reporting of growing coverage of genuine ethical concerns with machine-learning-driven AI. In the book, Smith posits narrative links between the technology and eugenics, slavery, and denial of women's suffrage. Most recently, coverage of the technology underpinning mobile-app-based Covid-19 contact tracing has strongly focused on themes of surveillance, authoritarianism, and social control. These are narratives that have become recurring themes in the discussions surrounding AI.

The growing work undertaken to make AI a more *trustworthy* technology is taking place in a challenging narrative context. Translating that work into a dividend of greater *trust* in the technology, given the narratives already at play, will take a concerted and skillful strategic communications effort.

The role of strategic communications

To achieve the wider AI objectives of innovation, widespread adoption, and standard-setting, states will have to develop strategies to speak to the public

23 See for example: 'UK Workers Who Lose Jobs to AI Will Be Retrained', BBC, 18 July 2019; Jonathan Vanian, 'Artificial Intelligence Will Obliterate These Jobs by 2030', *Fortune*, 19 November 2019; 'Workplace Automation, How AI Is Coming For Your Job', FT, September 2019.

about AI and data-driven technology, and about the work already being done to make these technologies worthy of public trust.

Strategic communications will need to address how we move away from a situation characterised by low levels of understanding of real AI, low levels of transparency about where and how it is being deployed, and general unease fuelled by a sense that society is not sufficiently in control of the technology.

Strategic communications encouraging greater trust in AI might seek to persuade the public to *think* that government is taking action to govern the technology, to *feel* that society is in the driving seat of how it develops, and to *share* data, so that the public will make use of AI services and support public sector innovation. To achieve these objectives, we will naturally need to communicate how the relevant technologies are being used and governed. We'll also need to demonstrate how government and civil society are working to make improvements to that governance. And we'll likely need to go even further, directly involving the public in the key debates about what needs to change.

Strategic communications seeking to build trust in AI is in its infancy. As it develops, a range of challenges will need to be overcome.

The historical background and sensational nature of the most common narratives about AI—killer robots, AI slave uprisings, mass unemployment—makes them dominant and difficult to shift. Especially when the conversations that we actually need to have with the public are about the more technical, difficult-to-grasp issues catalogued in *Rage*. In their earlier work, the editors of *AI Narratives* identified this problem.[24]

Those state actors most motivated to build trust in AI are likely to be at a credibility disadvantage with the very audiences they seek to influence. The polling referred to above showed that just 30% in the UK trust the government to use their data ethically; and when some of the most important ethical issues relate to increasing state power, skepticism about government intentions is natural. For similar reasons, the relationship between governments on the one hand, and the campaign groups and civil society organisations that drive much of the public conversation on the other, is a complicated one.

24 '[A]n over-emphasis on implausible AI and humanoid robotics could overshadow issues that are already creating challenges today. These issues are often harder to describe through compelling narratives.' *Portrayals and Perceptions of AI and Why They Matter*, The Royal Society, November 2019.

A long-term approach will need to generate a consistent positive trend in understanding and supporting the use of AI in the face of the inevitable scandals and setbacks to come. Given that public understanding of the technology and the ethical challenges it brings is low, greater regulatory enforcement and public debate may cause trust in AI to *fall* in the short-term.

Public participation in the complex ethical discussions that AI has generated is seen as an important component of increasing trust. In the UK, the Data Protection Regulator[25] and the Centre for Data Ethics and Innovation[26] have both carried out public engagement work on AI ethics issues, mirrored by a wide range of similar activity in civil society.[27]

Bringing together representative members of the public to meet in person, using expert witnesses to help them understand how AI works, and facilitating discussions on the ethical implications, is expensive. A relatively focused series of dialogue sessions with 100 members of the public is likely to cost upwards of £100,000. This expense means that statistically significant numbers of the general public cannot be directly involved in public dialogue activities. Creative approaches will be needed to leverage smaller projects that create a *sense* of widespread, even national, participation in the debate.

Perhaps most importantly, we lack comprehensive evidence on the current state of the AI communications space. Well-designed strategic communications require a solid evidence base about audiences, narratives and influencers. Communications aimed at promoting trust in AI will, therefore, necessitate a good understanding of narratives audience segments, and influencers in that specific space. Understanding the communications space is the most urgent next step for governments and civil society because action must be taken—but how communications will land is hard to judge, and the costs of getting it wrong are high.

AI Narratives is an excellent starting point for the most influential historical and literary narratives on AI. But we need to understand how these are actually playing out among the public and, especially, how they are interacting with our growing understanding of what machine-learning-based AI can actually do.

25 'Artificial Intelligence Citizens' Juries', UK Information Commissioner's Office, 18 February 2019.
26 'CDEI Concludes a Programme of Public Engagement Workshops on the Ethics of Online Targeting', UK Centre for Data Ethics and Innovation, 31 July 2019.
27 *How To Stimulate Effective Public Debate on the Ethics of Artificial Intelligence*, Involve, 2019.

DOI 10.30966/2018.RIGA.8.6.

The final chapter of *AI Narratives* details what can be done with AI itself to gather information on how AI is being discussed online. However, work like this will need to be complemented by traditional quantitative and qualitative audience and media research, including long-term consistent measurement of understanding of and trust in AI.

Conclusion

We must remain open-minded about where that research leads us. The dominant historical themes explored by *AI Narratives* might be easy to plug into, given the space they already occupy in the public imagination. Equally, they might be so far removed from the technology as it really is that they prove a distraction. Only innovative qualitative research, experimentation, and iteration can help us feel our way towards a resolution of this dilemma. That process may demonstrate that existing narratives can be repurposed in some way; it may prove that we are better off talking about very specific types and implementations of data-driven technology.

The term 'AI' itself, with its many meanings, might be an unhelpful one, too weighed down with historical baggage to be useful moving forward. *AI Narratives* is a useful exposition of the sheer weight of that baggage, which has so little to do with the technology that governments want to promote. 'Data-driven technology', 'algorithmic decision-making', and 'machine learning' are just some of the alternative terms in circulation, but entirely new ones may be needed. We will need to be flexible about how best to talk about these technologies—and our work to govern them—in a way that connects with the people whose trust we seek.

The Covid-19 pandemic has thrown into sharp relief (i) the important benefits AI can offer, (ii) why trust is crucial to realising those benefits, and (iii) the role that narratives—especially those around major events—play in determining levels of trust. Put simply, it has never been more important for governments to build and maintain the trust of their citizens in data-driven technology. Strategic communications will be a vital tool in this, but there are significant pitfalls that must be avoided. Solid research is needed to help us tell new stories about a surprisingly old idea in ways that serve our AI present and future.