Rethinking AI for Good Governance

Helen Margetts

This essay examines what AI can do for government, specifically through three generic tools at the heart of governance: detection, prediction, and data-driven decision-making. Public sector functions, such as resource allocation and the protection of rights, are more normatively loaded than those of firms, and AI poses greater ethical challenges than earlier generations of digital technology, threatening transparency, fairness, and accountability. The essay discusses how AI might be developed specifically for government, with a public digital ethos to protect these values. Three moves that could maximize the transformative possibilities for a distinctively public sector AI are the development of government capacity to foster innovation through AI; the building of integrated and generalized models for policy-making; and the detection and tackling of structural inequalities. Combined, these developments could offer a model of data-intensive government that is more efficient, ethical, fair, prescient, and resilient than ever before in administrative history.

From the 2010s onward, data-fueled growth in the development of artificial intelligence has made tremendous leaps forward in scientific advancements, medical research, and economic innovation. AI research and development is generally carried out by or geared toward the private sector, rather than government innovation, public service delivery, or policy-making. However, governments across the world have demonstrated strong interest in the potential of AI, a welcome development after their disinterested approach to earlier digital systems.[1] Security, intelligence, and defense agencies tend to be the most advanced, but AI is starting to be used across civilian policy sectors, at all levels of government, to tackle public good issues.[2]

What would a public sector AI look like? What might it offer to government in terms of improving the delivery of public goods and the design of policy interventions, or in tackling challenges that are specific to the public sector? Using a broad definition of AI that includes machine learning (ML) and agent computing, this essay considers the governmental tasks for which AI has already proved helpful: detection, prediction, and simulation. The use of AI for these generic governmental tasks has both revealed and reinforced some key ethical requirements of fairness, transparency, and accountability that a public sector AI would need to meet with new frameworks for responsible innovation. The essay goes on to discuss

where the development of a distinctively public AI might allow a more transformative model for government: specifically, developing internal capacity and expertise, building generalized models for policy-making, and, finally, going beyond the development of ethical frameworks and guidance to tackle long-standing inequalities and make government more ethical and responsive than it has ever been before.

Computers were first adopted by the largest departments of the largest governments in the 1950s.[3] In the very early days, government was an innovator and leader in digital technologies: the UK Post Office produced the world's first digital programmable computer in 1943, later used for code-breaking at Bletchley Park.[4] But since then, in many or even most countries, governments' digital systems were progressively outsourced, often in very large contracts that stripped digital expertise from the government. Partly for that reason, governments were slow to adopt Internet-based services or communicate with citizens online; in general (there are exceptions), they have lagged behind the private sector in adopting the latest generation of data-intensive technologies.[5] However, there has recently been much greater interest in the possibilities of data science and AI for government. The number of UK government announcements that mentioned data science and artificial intelligence rose from fifteen in 2015 to 272 in 2018. In the United States, a comprehensive study of the use of AI in the federal government found that nearly half of federal agencies studied (45 percent) had experimented with AI and related machine learning tools by 2020.[6] AI has helped governments perform three key tasks: detection, prediction, and simulation, all of which can improve policy-making and service delivery.[7] In a perhaps unanticipated way, AI also forces governments to think about ethical issues and the ethos of the government's digital estate, often in ways that have not been explicitly discussed before.

Governments need *detectors*: instruments for taking in information. Detection is one of the "essential capabilities that any system of control must possess at the point where it comes into contact with the world outside,"[8] and governments are no exception. They need to understand societal and economic behavior, trends, and patterns and calibrate public policy accordingly. To do this, governments need to detect (and then minimize) unwanted behavior by firms or individual citizens. For example, regulators need to be able to detect harmful behavior in digital environments, where the machine learning capabilities of large firms challenge traditional regulatory strategies and where the countering of online harms requires constant innovation.

Machine learning's core competency in classification and clustering offers government new capability in the detection and measurement of unwanted activity in large data sets. For example, machine learning is valuable in the detection of online harms such as hate speech, financial scams, problem gambling, bully-

ing, misleading advertising, or extreme threats and cyberattacks. Many agencies or regulators either need to detect these harms, or to oversee firms in so doing, requiring the building of machine learning "classifiers" trained on data generated by social media or other digital platforms. Growth of what is broadly called "counter-adversarial technology" to counter online threats to state or society is a particularly important task for "public" AI research and development, requiring constant innovation, as offenders continually game platforms to evade detection.[9] These techniques are of increasing importance to security and intelligence agencies, going beyond the creation of dedicated red teams for adversarial testing[10] to the creation of generative adversarial networks (GANs), in which neural networks are designed in tandem: one designed to be a generative network (the forger) and the other a discriminative network (the forgery detector). Each network can "train and better itself off the other, reducing the need for big labelled training data."[11]

Civilian agencies across sectors also benefit from enhanced detection capabilities. For example, the U.S. Securities and Exchange Commission uses a historical data set of past issuer filings and machine learning with a random forest model to identify which filers might be engaged in suspect earnings management, relying on indicators such as earnings restatements and past enforcement actions.[12] Detection is enhanced by AI-powered developments in robotics, computer vision, and spatial computing. Health research agencies have been particularly advanced in the use of computer vision and machine learning models trained to detect early signs of, for example, cancer. Law enforcement agencies have been early adopters of AI for detection, combining these tools with robotic devices and AI-related technologies such as computer vision. The U.S Department of Homeland Security's Customs and Border Protection (CBP) agency has a long running program of using facial recognition technology, growing out of the agency's emphasis on counterterrorism post 9/11, developed by a range of private vendors using deep learning within their proprietary technologies.[13]

The *predictive capacity* of machine learning has much to offer regulatory agencies and governments broadly, which are not known for their strength in foresight or forecasting. Governments can use machine learning tools to spot trends and relationships that might be of concern or identify failing institutions or administrative units. For example, in 2020, the U.S. Food and Drug Administration used machine learning techniques to model relationships between drugs and hepatic liver failure, with decision trees and simple neural networks used to predict serious drug-related adverse outcomes. They utilized regularized regression models, random forests, and support vector techniques to construct a rank-ordering of reports based on their probability of containing policy-relevant information about safety concerns, allowing the agency to prioritize those most likely

to reveal problems.[14] More generally, the use of predictive risk-based models can greatly enhance the prioritization of sites for inspection or monitoring, from water pipes, factories, and restaurants to schools and hospitals, where early signs of failing organizations or worrying social trends may be picked up in transactional data.

Government agencies can use AI tools to predict aggregate demand, for example, in schools, prisons, or children's care facilities. Understanding future needs is valuable for resource planning and optimization, allowing government agencies to direct human attention or manpower where it is most required. Machine learning models of COVID-19 spread during 2020–2021 might have been used to direct resources such as ventilators, nurses, and drug treatments toward those areas likely to be most affected, and even to target vaccination programs. An investigation of data science in UK local government suggested that even in 2018, 15 percent of local authorities in the United Kingdom were using data science to build some kind of predictive capability, such as to target safety prevention measures at the streets placing most demand on emergency services.[15] Unsupervised learning models are also utilized to categorize criminal activities from free-text data generated by complaints, of potential use across the UK criminal justice system.[16]

The use of prediction to deliver individual (as opposed to aggregate) risk scores is much more controversial. For local authorities that have used predictive techniques to identify the number of children that are likely to be at risk of abuse or neglect, the next step from forecasting (say) demand for childcare places is likely to be "which children?" Such a question would come naturally to social services departments terrified of being held responsible for the next ghastly case of abuse to hit the headlines, the next "Baby P." But should a technique that is essentially inductive be used in this way? A risk of 95 percent of being a victim of an abusive incident means that there is still a chance that the event will not happen, and if the figure is 65 percent, the meaning of the individual number is highly ambiguous. Social policy experts who advocate this kind of machine learning for decision support have built models to support childcare workers' decision-making in New Zealand, the United States, and Australia.[17] But other studies have counseled a more cautious and thoughtful approach, and noted the importance of the data environment.[18] The most feted version, in Pittsburgh, was built from a data-rich environment providing a 360-degree view of all children's and their families' interactions with state agencies throughout their lives, an environment that rarely exists in local authorities. And such systems are extremely vulnerable to bias, especially where data are derived from the criminal justice system.

As with detection, the earliest examples of the use of machine learning for risk prediction came from law enforcement agencies. In the United States, a prominent example was the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system, a decision support system for judges that assesses the risk of an individual prisoner being likely to reoffend, and therefore in-

forming sentencing decisions. The judges receive risk scores in low, medium, and high risk buckets, and feed this evidence into the decision-making process. A 2016 study by ProPublica showed that COMPAS exhibited racial bias, a claim that has generated much discussion over this use of machine learning in legal judgments.[19] The system also demonstrates some of the subtle but deep shifts in perceptions within the policy-making system that occur when machine learning technologies are introduced, bringing with them notions of statistical prediction to a "situation which was dominated by fundamental uncertainty about the outcome before," according to one thoughtful case study on the implementation of COMPAS, showing that practitioners within the system valued what they perceived as the "research-based" nature of COMPAS results, which they felt reduced uncertainty in the system.[20]

The third area in which AI-related technologies can help policy-makers in the design of policy interventions and evidence-driven, data-intensive decision-making is *simulation*. Governments need ways of testing out interventions before they are implemented to understand their likely effects, especially those of costly new initiatives, major shifts in resource allocation, or cost-cutting regimes aimed at saving public resources. In the past, the only option for trying out initiatives was by running field experiments: randomized trials in which the intervention is applied to a "treatment group" and the results are compared with a "control group." But such trials are expensive and take a long time, challenge notions of public equity, and sometimes are just not possible due to attrition or ethical constraints.[21] In contrast, the availability of large-scale transactional data, and innovative combinations of agent computing and machine learning, allow the simulation of interventions so unintended consequences can be explored without causing harm.

Like AI itself, agent computing is a form of modeling that has been in existence for a long time but has been revolutionized by large quantities of data. The agent-based method was developed within economics in the 1960s and 1970s for the purposes of simulation, but these were "toy models": formal models with hardly any data, and when tested on data generated by real-life situations, they tended to perform very badly indeed. In contrast, the kind of agent computing models used now are based on large-scale data, which can replicate whole economies, with 120 million firms and workers.[22] A modern agent-based model like this consists of individual software agents, with states and rules of behavior, and large corpuses of data pertaining to the agents' behavior and relationships. Some computer scientists have called for such models to be developed *ex ante* – "agent-based modeling as a service" – so that in an emergency, it could be rapidly employed to feed in key variables and model possible policy interventions. Mainstream economics has been resistant to such innovations, and political systems have inbuilt tenden-

cies to try out hurried policy decisions, such as not having enough police, or doctors or nurses, and learning the hard way. But the disadvantages of this on-the-hoof policy-making were illustrated during the first stage of the COVID-19 crisis in 2020, when in many countries, policies regarding masks, social distancing, and lockdown measures were made in an ad hoc and politically motivated fashion.

Agent computing has gradually gained popularity as a standard tool for transport planning, or to provide insight for decision-makers in disaster scenarios such as a nuclear attack or pandemic.[23] Researchers working with police forces are trialing the use of large-scale, real-time transactional data from daily activities of individual police in an agent-based model that would allow police managers to try out different levels of police resourcing and measure the potential effects on delivery of criminal justice.[24] If viable, such models could have potential for other areas of the public sector, where large quantities of trained professionals are needed, such as in education or health care. In this way, agent computing can be another good way of optimizing resources, by testing out the impact of different levels of manpower without experiencing unintended consequences. Similarly, the United Nations Development Programme is using an agent computing model to help developing countries work out which policies – such as health, education, transportation, and so on – should be prioritized in order to meet their sustainable development goals.[25] Researchers have started to explore the possibilities of "societal digital twins": a combination of spatial computing, agent-based models, and "digital twins," or virtual data-driven replicas of real-world systems. These have become popular for physical systems in engineering or infrastructure planning, although proponents warn that the complexity of social systems renders the social equivalent of digital twins "a long way from being able to simulate real human systems."[26]

G overnments of the progressive era of public administration from the late nineteenth and early twentieth centuries stressed the need for a "public service ethos" to limit corruption, waste, and incompetence. Such an ethos prioritized values of honesty and fairness in an attempt to distinguish public officials from the "inherently venal" nature of politicians and an increasingly corrupt private sector.[27] But as state operations became increasingly automated, and personnel were replaced with digital systems, which were then outsourced to computer services providers, there was a diminishing sense in which this ethos could be said to apply to government's digital estate.[28] The advent of AI, however, has forced a rethink about the need to address issues of fairness, accountability, and transparency in the way that government uses technology, given that they pose greater challenges to these values than earlier generations of technology used by government.

It is around ethical questions such as fairness that the distinctiveness of the public sector becomes stark. If (say) Amazon uses sophisticated AI algorithms to

target customers in a biased way, it can cause offense, but it is not on the same scale as a biased decision over someone's prison sentence or benefits application. Users of digital platforms know very little about the operation of search or news-feed algorithms, yet will rightly have quite different expectations of their right to understand how decisions on their benefit entitlement or health care coverage have been made. The opaqueness of AI technology is accepted in the private sector, but it challenges government transparency.

From the late 2010s onward, there has been a burgeoning array of papers, reviews, and frameworks aimed at tackling these issues for the use of AI in the public sector. The most comprehensive and widely used across the UK government is based on the principles of fairness, accountability, trustworthiness, and transparency, and a related framework was applied to the use of AI in the COVID-19 crisis.[29] Policy-makers are starting to coalesce around frameworks like these, and ethics researchers are starting to build the kinds of tools that can make them usable and bring them directly into practice. It might be argued that progress is greater here than it has been in the private sector. There is more willingness to contemplate using less innovative – or *differently* innovative – models in order, for example, to make AI more transparent and explainable in the process of high-stakes decisions or heavily regulated sectors.[30]

The development of such frameworks could lead to a kind of public ethos for AI, to embed values in the technological systems that have replaced so much of government administration. Such an ethos would not just apply to AI, but to the legacy systems and other technologies that first started to enter government in the 1950s, and could be highly beneficial to the public acceptance of AI.[31] There is a tendency to believe that the technological tide will wash over us, fueled by media and business school hype over "superintelligent" robots and literary and cinematic tropes of robots indistinguishable from humans, powered by general AI. If we do not design appropriate accountability frameworks, then politicians and policy-makers will take advantage of this blame-shifting possibility. This will range from cases like the UK prime minister blaming poor statistical processes to calculate public examination results after school closures in the 2020 pandemic prevented exams from taking place as a "mutant algorithm," to the more nuanced and unconscious shifting of responsibility to statistical processes involved in judicial decision-making with AI observed above. A public sector AI in which fairness, accountability, and transparency are prioritized would be viewed as more trustworthy, working against such perceptions.

S o in what areas might government do more with AI? By 2021, government's use of AI was starting to speed up; the large-scale study of the use of AI by the U.S. federal government concluded in 2020 that "though the sophistication of many of these tools lags behind the private sector, the pace of AI/ML de-

velopment in government seems to be accelerating."[32] However, there are various ways that AI could have a more transformative effect.

First, governments could prioritize the development of expertise and capacity in AI to foster innovation and overcome some of the recurring challenges. As noted above, the history of government computing has been characterized by large-scale contracting to global computer services providers, but AI does not lend itself to this kind of outsourcing, whereby governments lose control of key features. For example, the U.S. CBP was criticized in 2020 for being unable to explain failure rates of biometric scanning technology "due to the proprietary technology being used."[33] Similar issues have dogged the adoption of facial recognition technologies by police agencies, with moratoria announced in several cities. There is evidence that government agencies realize the importance of developing capacity: the same U.S. study also found that "over half of applications were built in-house, suggesting there is substantial creative appetite within agencies."

An area with great scope is the use of data-intensive technologies to develop new generalized models of policy-making. Governments have little tradition of using transactional data to inform decision-making. In the classic Weberian model of bureaucracy, data are compressed within files, available for checking individual pieces of information, but generating no usable data for analytics.[34] This characteristic of governments' information architecture persisted into the era of computerization, with a lack of usable data remaining a feature of the "legacy systems" of many governments. This point was well illustrated during the first wave of the COVID-19 pandemic, when many countries discovered that they lacked the kinds of data and modeling that could help design interventions. Key data flows did not exist in real time; in the United Kingdom, for example, it turned out that data for deaths were available only several weeks after the death had occurred. Data were not fine-grained enough; the design of a stimulus package requires sectoral-level data in order to target resources to those firms most in need. Modeling took place in silos such as public health, health care, education, or the economy, meaning that interventions were targeted only at (say) economic recovery or the health crisis, rather than an integrated approach taking account of the fact that the domains were intertwined. Resilient policy-making would involve building such data flows and using agent computing, machine learning, and other AI methodologies to create integrative models to both recover from the current crisis and face future shocks.[35]

Finally, perhaps the most ambitious use of AI would be to tackle issues of equality and fairness in governmental systems in a profound and transformative way, identifying and reforming long-standing biases in resource allocation, decision-making, the administering of justice, and the delivery of services. Many of the causes of bias and unfairness in machine learning, for example, come from training data generated by the existing system. The COVID-19 pandemic revealed

many structural inequalities in how citizens are treated – for example, in the delivery of health care to people from different ethnic groups – just as the mobilization around race has revealed systemic racism in police practice. Data and modeling have made these biases and inequalities explicit, sometimes for the first time. Some researchers have suggested that we might develop AI models that incorporate these different sources of data and combine insights from a range of models (so-called ensemble learning) aimed at the needs of different societal groups.[36] Such models might be used to produce unbiased resource allocation methods and decision support systems for public professionals, helping to make government better, in every sense of the word, than ever before.

Artificial intelligence can help with core tasks of government. These technologies can enable real-time, transactional data to enhance government's armory of detecting tools, to build predictive models to support decision-making, and to use simulation to design policy interventions that avoid unintended consequences. They face distinct ethical challenges when used for these public sector tasks, requiring new frameworks for responsible innovation. As policy-makers become more sophisticated in their use of AI, these technologies might be developed to overcome fragilities exposed in the COVID-19 pandemic, to create new, more resilient models of policy-making to face future shocks, and to "build back better," the catchphrase of many governments in the postpandemic era. AI can reveal and perhaps mitigate some structural biases and might even be used to tackle some profound inequalities in the distribution of resources and the design and the delivery of public services such as education and health care. This would require a specific branch of AI research and development, geared at distinctively public sector tasks and needs. Such a remit would be no less complex or challenging than for any other field of AI. Indeed, some deep learning experts suggest that even where machine learning has had success, as in medical diagnosis of X-ray images, models are still outperformed by human radiologists in clinical settings.[37] But the potential public good benefits are huge.

ABOUT THE AUTHOR

Helen Margetts OBE FBA is Professor of Society and the Internet at the Oxford Internet Institute at the University of Oxford and Programme Director for Public Policy at The Alan Turing Institute, London. She is the author of, among other publications, *Political Turbulence: How Social Media Shape Collective Action* (with Peter John, Scott Hale, and Taha Yasseri, 2016), *Digital Era Governance: IT Corporations, the State, and E-Government*, 2nd ed. (with Patrick Dunleavy, Simon Bastow, and Jane Tinkler,

2008), and *The Tools of Government in the Digital Age* (with Christopher Hood, 2007), and editor of *Paradoxes of Modernization: Unintended Consequences of Public Policy Reform* (with Perri 6 and Christopher Hood, 2010).

ENDNOTES

[1] Helen Margetts, *Information Technology in Government: Britain and America* (New York: Routledge, 1999).

[2] Thomas M. Vogl, Cathrine Seidelin, Bharath Ganesh, and Jonathan Bright, "Smart Technology and the Emergence of Algorithmic Bureaucracy: Artificial Intelligence in UK Local Authorities," *Public Administration Review* 80 (6) (2020): 946–961, https://onlinelibrary.wiley.com/doi/10.1111/puar.13286.

[3] Margetts, *Information Technology in Government*.

[4] Mar Hicks, *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing* (Cambridge, Mass.: MIT Press, 2017).

[5] Margetts, *Information Technology in Government*; Patrick Dunleavy, Helen Margetts, Jane Tinkler, and Simon Bastow, *Digital Era Governance: IT Corporations, the State, and E-Government* (Oxford: Oxford University Press, 2006); and Helen Margetts and Cosmina Dorobantu, "Rethink Government with AI," comment, *Nature*, April 9, 2019, 163–165.

[6] David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey, and Mariano-Florentino Cuéllar, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies* (Washington, D.C.: Administrative Conference of the United States, 2020).

[7] Margetts and Dorobantu, "Rethink Government with AI."

[8] Christopher Hood and Helen Margetts, *The Tools of Government in the Digital Age* (London: Macmillan, 2007), 3.

[9] Abhijnan Rej, "Artificial Intelligence for the Indo-Pacific: A Blueprint for 2030," *The Diplomat*, November 27, 2020; and Bertie Vidgen, Alex Harris, Dong Nguyen, et al., "Challenges and Frontiers in Abusive Content Detection," *Proceedings of the Third Workshop on Abusive Language Online*, Florence, Italy, August 1, 2019.

[10] National Security Commission on Artificial Intelligence, *Final Report* (Washington, D.C.: National Security Commission on Artificial Intelligence, 2021), 383.

[11] Ibid., 607; and Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, et al., "Generative Adversarial Nets," *Neural Information Processing Systems* 27 (2014), https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf.

[12] Engstrom et al., *Government by Algorithm*, 23.

[13] Ibid., 32.

[14] Ibid.

[15] Jonathan Bright, Bharath Ganesh, Cathrine Seidelin, and Thomas M. Vogl, "Data Science for Local Government" (Oxford: Oxford Internet Institute, University of Oxford, 2019); and Vogl et al., "Smart Technology and the Emergence of Algorithmic Bureaucracy."

16 Daniel Birks, Alex Coleman, and David Jackson, "Unsupervised Identification of Crime Problems from Police Free-Text Data," *Crime Science* 9 (1) (2020): 1–19.

17 Rhema Vaithianathan, Emily Putnam-Hornstein, Nan Jiang, et al., *Developing Predictive Models to Support Child Maltreatment Hotline Screening Decisions: Allegheny County Methodology and Implementation* (New Zealand: Centre for Social Data Analytics, AUT University, 2017); and Rhema Vaithianathan, "Five Lessons for Implementing Predictive Analytics in Child Welfare," *The Chronicle of Social Change*, August 29, 2017, https://chronicleofsocialchange.org/opinion/five-lessonsimplementing-predictive-analytics-child-welfare/27870.

18 David Leslie, Lisa Holmes, Christina Hitrova, and Ellie Ott, *Ethics Review of Machine Learning in Children's Social Care* (London: What Works for Children's Social Care, 2020), https://www.turing.ac.uk/research/publications/ethics-machine-learning-childrens-social-care.

19 Alex Chohlas-Wood, "Understanding Risk Assessment Instruments in Criminal Justice," Brookings Institution's series on AI and Bias, June 19, 2020, https://www.brookings.edu/research/understanding-risk-assessment-instruments-in-criminal-justice/.

20 See Aleš Završnik, "Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings," *European Journal of Criminology* 18 (5) (2019).

21 Helen Margetts and Gerry Stoker, "The Experimental Method," in *Theory and Methods in Political Science*, ed. Vivien Lowndes, David Marsh, and Gerry Stoker (London: Macmillan International Higher Education, 2017); and Peter John, *Field Experiments in Political Science and Public Policy: Practical Lessons in Design and Delivery* (New York: Routledge, 2019).

22 Robert Axtell, "Endogenous Firm Dynamics and Labor Flows via Heterogeneous Agents," in *Handbook of Computational Economics*, 4th ed., ed. Cars Hommes and Blake LeBaron (Amsterdam: North-Holland, 2018), 157–213.

23 M. Mitchell Waldrop, "Free Agents," *Science* 360 (6385) (2018): 144–147.

24 Julian Laufs, Kate Bowers, Daniel Birks, and Shane D. Johnson, "Understanding the Concept of 'Demand' in Policing: A Scoping Review and Resulting Implications for Demand Management," *Policing and Society* 31 (8) (2020): 1–24.

25 Omar A. Guerrero and Gonzalo Castañeda, "Policy Priority Inference: A Computational Framework to Analyze the Allocation of Resources for the Sustainable Development Goals," *Data & Policy* 2 (2020).

26 Dan Birks, Alison Heppenstall, and Nick Malleson, "Towards the Development of Societal Twins," *Frontiers in Artificial Intelligence and Applications* 325 (2020): 2883–2884.

27 Christopher Hood, "A Public Management for All Seasons?" *Public Administration* 69 (1) (1991): 3–19; and Christopher Hood, *Explaining Economic Policy Reversals* (Buckingham: Open University Press, 1994).

28 Margetts, *Information Technology in Government*; and Dunleavy et al., *Digital Era Governance*.

29 Engstrom et al., *Government by Algorithm*; and David Leslie, "Tackling COVID-19 through Responsible AI Innovation: Five Steps in the Right Direction," *Harvard Data Science Review* (2020).

30 United Kingdom Information Commissioner's Office and The Alan Turing Institute, *Explaining Decisions Made with AI* (Wilmslow, United Kingdom: Information Commissioner's Office, 2020), https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-ai/.

31 Helen Margetts, "Post Office Scandal Reveals a Hidden World of Outsourced IT the Government Trusts but Does Not Understand," *The Conversation*, April 29, 2021.

32 Engstrom et al., *Government by Algorithm*, 55 ; and Vogl et al., "Smart Technology and the Emergence of Algorithmic Bureaucracy."

33 Engstrom et al., *Government by Algorithm*, 33–34.

34 Patrick Dunleavy and Helen Margetts, *Digital Era Governance and Bureaucratic Change* (Oxford : Oxford University Press, 2022).

35 Jessica Flack and Melanie Mitchell, "Uncertain Times," *Aeon*, August 21, 2020, https:// aeon.co/essays/complex-systems-science-allows-us-to-see-new-paths-forward ; and B. MacArthur, Cosmina Dorobantu, and Helen Margetts, "Resilient Policy-Making Requires Data Science Reform," *Nature* (under review).

36 MacArthur et al., "Resilient Policy-Making Requires Data Science Reform."

37 Tekla S. Perry, "Andrew Ng X-Rays the AI Hype," IEEE Spectrum, May 2, 2021, https:// spectrum.ieee.org/view-from-the-valley/artificial-intelligence/machine-learning/ andrew-ng-xrays-the-ai-hype.