

Socializing Data

Diane Coyle

Will the proliferation of data enable AI to deliver progress? An ever-growing swath of life is available as digitally captured and stored data records. Effective government, business management, and even personal life are increasingly suggested to be a matter of using AI to interpret and act on the data. This optimism should be tempered with caution. Data cannot capture much of the richness of life, and while AI has great potential for beneficial uses, its delivery of progress in any human sense will depend on not using all the data that can be collected. Moreover, the more digital technology rewires society, creating opportunities for the use of big data and AI, the greater the need for trust and human deliberation.

Data have always been important for government and policy. Statistics are, as the name suggests, categorized data useful for states.¹ States have collected and collated data for centuries, not least for the purposes of taxation. Censuses too are ancient, defining the boundaries of power, though they are likely to be replaced by other government-collected data sets about individuals.² The purpose of governmental measurement is to create conceptual order, to classify the vast array of possible data points into meaningful categories, enabling better decisions. Over the quarter-millennium of modern economic growth, the scope of data collection and processing into statistics has become increasingly extensive.

In *Seeing like a State* (1998), political scientist James Scott argues that modern states classify reality to improve the *legibility* of what they govern, to better control it. He writes: “Legibility implies a viewer whose place is central and whose vision is synoptic. . . . This privileged vantage point is typical of all institutional settings where command and control of complex human activities is paramount.”³ Many of his examples of states bending reality into order concern economic activities such as forestry or agriculture, with reality conforming increasingly to the classifications devised to understand it. There is a feedback loop whereby statistics collect and classify data points found in the wild, then subsequently influence activities and shape reality over time, so that future data will be more likely to fit into the predefined categories. This has been described by statistician André Vanoli as “the dialectic of appearance and reality.”⁴ Or as historian Theodore Porter put it, “The quantitative technologies used to investigate social and economic life always work best if the world they aim to describe can be remade in their image.”⁵

For example, the principal measure of economic progress since the early 1940s has been gross domestic product (GDP).⁶ Governments gear their policies toward increasing GDP, and people duly respond to the incentives created by policies such as tax breaks, subsidies, public infrastructure investment, or cheaper meals out.⁷ Disappointing statistics can topple governments, as they did with the UK Labour government of the late 1970s, paving the way for the Thatcherite revolution. GDP has not been a terrible metric for progress: compared with previous generations, our living standards are without doubt higher. We have better health, more leisure, more comfortable homes, and the convenience of many new technologies. Yet even at the dawn of GDP's invention, some realities had to be bent to fit the statistical framework. Some were rendered invisible, defined as being outside "the economy," such as household work and nature. Without nature, there is no economy and yet the consequences for sustainability of this fateful definitional choice are becoming all too clear, and the progress we thought we had is at least partly illusory.

Reality and the statistical picture also diverge when reality is changing. As statistician Alain Desrosières has written, "In its innovative phase, industry rebels against statistics because, by definition, innovation distinguishes, differentiates and combines resources in an unexpected way. Faced with these 'anomalies,' the statistician does not know what to do."⁸ At present, for official statisticians, life is one damned anomaly after another. For just as agriculture's share was overtaken by manufacturing in the industrial revolution, the material economy is smaller now relative to the dematerializing economy of digitally enabled services.⁹ The statistical categories no longer fit well. Paradoxically, in the economy of ever more data, it is proving increasingly difficult to bring informational order, for the state to gain that desired legibility.

This is a paradox because the promise of big data and its use in AI has inspired renewed visions inside government of enhanced legibility. Such visions are not new. From the late 1950s onward, computers have seemed to promise a clearer, synoptic understanding of society.¹⁰ One ambitious 1970s project was Project Cybersyn in Salvador Allende's Chile, administered by cyberneticist Stafford Beer, which was intended to implement an efficiently planned economy.¹¹ A similar vision of data-enabled, improved legibility has revived in the big data digital era. On the left of UK politics it found expression as "fully automated luxury communism."¹² In the UK Conservative government elected in 2019, it took physical shape as a control room at the heart of government, and a UK Strategic Command contract with tech firm Improbable to build a "digital twin," a simulation of the whole of Britain.¹³ The fact that both ends of the political spectrum envision data-driven efficiency suggests a big data rerun of the 1930s socialist calculation debate.¹⁴

The thing that is seen in seeing rooms of these kinds – physical rooms with displays of information to inform decision-makers – is ordered data. There is a kind of commodity fetishism regarding the mechanics of displaying the data. The technology of data has long been glamorous, arousing intense public and political interest. The great exhibitions and world’s fairs of the nineteenth and early twentieth centuries had popular displays of high-tech data management artifacts such as filing cabinets and cards.¹⁵ The same is true of digital technology and Silicon Valley, which have inspired numerous nonfictional and fictional accounts. Databases have changed form over time as physical hardware and computational power have evolved, so the embodiment and the usability (searchability) of data have not been constant, and the technologies of display combined with the classification and conceptual framework organizing the data affect the way decision-makers understand the world. The emphasis on the synoptic view – through a computer simulation, through a room kitted out with the latest screens and data feeds – is an assertion of political control through greater legibility. Then-UK government adviser Dominic Cummings presented it as a matter of public interest:

There is very powerful feedback between: a) creating dynamic tools to see complex systems deeper (to see inside, see across time, and see across possibilities), thus making it easier to work with reliable knowledge and interactive quantitative models, semi-automating error-correction etc, and b) the potential for big improvements in the performance of political and government decision-making.¹⁶

In other words, the claim is that data science and AI, suitably embodied in a seeing room, can be the vehicle for delivering “high performance” by government.

However, the emphasis is on the technologies of cognition and management, rather than the construction of the data going into the process, or the assessment of what constitutes improvement. The implicit assumption is that this is a determination made by the center, by those in the seeing room. This assumption is exactly why an ambition to use data for progress can embed biases, create ambiguity about accountabilities, or appear to be part of the surveillance society.¹⁷ There is certainly nothing new about state attempts to exercise comprehensive surveillance. East Germany’s Stasi offers an extreme recent example. Its data took analog form with a technological infrastructure turning data into seeing: card records with a bespoke filing cabinet technology, photographs, steam irons for opening mail, tape recordings, and computers. Despite the existence of formal regulations controlling access to this data, a citizen of the former German Democratic Republic was a *gläserne Mensch*, a transparent being. Perhaps we are all becoming transparent now. Digital technology makes the amassing of data records trivially cheap and easy by comparison with the 1980s, and security agencies have been doing this at scale.

Big tech companies, not just security agencies, have been amassing the biggest and best databases and the know-how to use them for a purpose. Their purpose is

profit, rather than public good, and their market power ensures they do not need to serve the interests of their users or the public in general. Big Tech's success vis-à-vis state power is amply evident in the erosion of national tax bases as ever more economic activity goes online. It is not clear how much governments can limit this.¹⁸ As being able to raise tax revenues is a core state function, there can be little question about the power of the biggest digital companies. If the synoptic view of what is happening anywhere is available to anybody, it is Google or Facebook. They, not officials or politicians, are collecting, categorizing, and using the new proliferation of data.

As long as data are seen as individual property amenable to normal market exchange, that will continue to be the case, despite recent regulatory moves in several jurisdictions to enforce some data sharing by the tech giants. The reason big tech companies have been able to acquire their power is the prevailing conceptual framework, crystallized into law, for understanding data as property. Rather than the appreciation that data reflect constructed categories, a particular lens or framework measuring and shaping reality, data are seen as the collection of natural objects: the classifications codified and programmed into data feeds just are what they are. These constructed data records are then subject to legal rules of ownership. Data are presumed to be transferred and owned by corporations as soon as the user of a service has accepted its terms and conditions.

The consequences of this property rights concept applied to data, or information, illuminate why it is so pernicious. For example, John Deere and General Motors (as corporate persons) have claimed in U.S. copyright courts that farmers or drivers who thought they were purchasing their vehicles do not in fact own them and have no right to repair them. The company's reasoning is that a tractor is no longer mainly a metal object whose ownership as a piece of property is transferred from John Deere to the farmer, but rather an intangible data-fed software service licensed from the company, which just happens to have a tractor attached.¹⁹ Indeed, screens with data about weather, soil conditions, and seed flow proliferate inside tractor cabins and feed into the diagnostic software installed by the manufacturer, which provides information to enable decisions raising crop yields. The John Deere claim to ownership of the intangible dominates the farmer's claim to ownership of the physical vehicle it is bundled with. To date, the courts have been largely sympathetic to the corporations and to the strong ownership claims made by Amazon over e-books, by makers of games on consoles, as well as by vehicle manufacturers.

One response to such corporate ownership over data and data processing claims has been the demand for corporations to pay for "data as labor."²⁰ With this, each data point an online business collects from users' activities would be rewarded with a small financial payment. However, as economist Zoë Hitzig and colleagues point out, this remedy also considers data as a transferrable, individual item of property, and implicitly as a natural object "given" by the underlying reality.²¹

The data-as-property perspective assumes data are an object in the world, with an independent reality, differing from other givens only in being intangible. Yet not only are data nonrival (their use does not deplete them so many people can use them), but they are also inherently relational. Data are social. Even when it comes to data that are seemingly ultrapersonal – for example, that I passed a particular facial recognition camera at a given moment – the information content and usefulness of the data are always relational.²² A facial image needs to be compared with a police database. Even then its utility for the purpose of detecting suspected criminals depends on the quality of the training data used to build the machine learning algorithm, including its biases, the product of a long history of unequal social relations. The relational character of data means they are both constructed by social relations and a collective resource for which market exchange will not be the best form of organization.²³ Indeed, this is why there are few markets for data; where data are sold – for example, by credit rating agencies – the market is generally thin, with no standardized, posted prices. The use value of data – their information content enabling decisions to be made – is highly heterogeneous.

That markets are a poor organizational model for the optimal societal use of data is Economics 101. Does that make government the right vehicle to use big data and AI for the public good? Can and should governments aim to beat big tech at the seeing game? The promise of automating policy through seeing rooms and use of AI is greater efficiency and, potentially, better outcomes. Yet there is increasing use of algorithmic processes in arenas in which decisions could have a large impact on people's lives, such as criminal justice or social security.

Much of the literature on the informational basis of organizations focuses on complexity as the constraint on effective information-processing, given an objective function.²⁴ Automation is superior in routine contexts: more reliable, more accurate, faster, and cheaper. What is more, machines deal more effectively with data complexity than humans do, given our cognitive limitations. This is a key advantage of machine learning systems as the data environment grows more complex. The system is better able than any human to discern patterns and statistical relationships in the data, and indeed the more complex the environment, the greater the AI advantage over human-scale methods. However, whenever there is uncertainty, the advantage tips back to humans. The more frequently the environment changes in unexpected ways, or the more dramatic the scale of change, the greater the benefits of applying human judgment. The statistical relationships on which automated decision rules are based will break down in such circumstances (in economics this is known as the Lucas critique).²⁵ The selection of a machine or human to make decisions is generally presented as a trade-off. However, it has long been argued, or hoped, that AI can improve the terms of this trade-off.²⁶

There are several reasons to doubt this. One is the well-known issue of bias in training data sets, the inevitable product of unfair societies in the way data are classified, constructed, collected, and ordered.²⁷ Any existing data set reflects both the classification framework used and the way that framework has shaped the underlying reality over time (that is, André Vanoli's dialectic referred to earlier). The data science community has become alert to this challenge and many researchers are actively working on overcoming the inevitable problems raised by data bias. But bias is not the only issue.

Another less well-recognized issue (at least in the policy world) is that decisions based on machine learning need an explicitly coded objective function. Yet in many areas of human decision-making – particularly the most sensitive, such as justice or welfare – objectives are often left deliberately implicit. Politics in democracies requires compromise on high-level issues so that low-level actions can be taken. These “incompletely theorized agreements” are not amenable to being encoded in machine learning (ML) systems, in which precision about the reward function is needed even if conflicting objectives are combined with different weights.²⁸ The further deployment of ML in applied policy practice may require more explicit statements of objectives or trade-offs, which will be challenging in any domain where people's views diverge.²⁹ There could be very many of these, even in policy areas that seem straightforward. For example, how should public housing be allocated? There has been a pendulum swinging over time between allocation based on need and allocation based on likelihood to pay rent. These are conflicting objectives, and yet many of the same families would be housed under either criterion.

The extensive discussions of value alignment in the AI literature tussle with how to combine the brutally consequentialist nature of AI with ambiguity or conflicts about values. Given any objective or reward function, ML systems will game their targets far more effectively than any bureaucrat ever did. All the critiques of target setting in the public management literature, on the basis that officials game these for their personal objectives, apply with extra force to systems automating target delivery. This has led to concerns – albeit overstated – about runaway outcomes far from what the human operators of the system wanted.³⁰ One possible avenue is inverse reinforcement learning – that is, when ML systems try to infer what they should optimize for – which can accommodate uncertainty about the objective, but takes the existing environment as the desired state of affairs.³¹ Political theorist and ethicist Iason Gabriel rightly emphasizes the need for legitimate societal processes to enable value alignment; but we do not have these yet.³²

Market arrangements based on the concept of private property transactions are inappropriate for data, given their relational characteristics. In economic terms, there are large externalities, whereby one individ-

ual's provision of data can have either negative (loss of privacy) or positive (useful information) implications for other people.³³ Rather than being considered as property amenable to market exchange, data instead need to be subject to governance arrangements of permitted access and use. Online, the offline norms of sociologist Georg Simmel's concept of "privacy in public" do not exist.³⁴ This concept refers to the norms people adopt limiting what they know about each other in their different roles. Even publicly available information (such as where somebody lives) is not made known in a specific context (such as the marking of an exam paper by their lecturer). These voluntary informational restraints and social relations of trust play an important role in sustaining desirable outcomes such as fairness, privacy, or self-esteem.³⁵ Similar norms do not exist online. Big tech joins up too many data about each of us. People can reasonably be concerned about government seeing rooms doing the same.

At the same time, some joining up of data for some uses could without question lead to improved outcomes for individuals. So we have ended up in the worst of all worlds: a "surveillance state" or "surveillance economy" in which valid privacy concerns about certain data uses prevent other uses of "personal" data for collective and individual good. Consider the successful argument that governments should not use data from COVID-19 apps to trace individuals' contacts during the pandemic, leading almost all governments to adopt the Google and Apple application programming interfaces (API) with privacy enforced, all the while as personal liberty was infringed through lockdowns tougher than would have been needed with effective contact tracing. Meanwhile, governments and researchers have been able to use big data and machine learning to inform policies during the pandemic but could have done much more to avert unequal health outcomes with linked data about individuals' health status, location, employment, ethnicity, and housing.

The debate about privacy has become overly focused on individual consent and data protection. It should be a debate about social norms and what is acceptable in different contexts, translated into rights of access and use for limited, specific purposes.³⁶ In both the commercial and the public sphere, the promise of AI for decision-making will not be realized unless the kind of information norms that operate offline are created online. The control of access and use is not just a technical issue but a social and political one.

As the world gets both more complex and more uncertain, big data and AI will need to socialize in another way, by combining with human judgment more often. The experiences of 2020, or the impact of extreme climate-related events from California burning to Texas freezing, are suggestive of the prospect that "radical uncertainty" will characterize the twenty-first century.³⁷ Anybody with any knowledge of forecasting (no matter how small or big the

data set) will know that uncertainty about future outcomes multiplies over time. “Further computational power doesn’t help you in this instance, because uncertainty dominates. Reducing model uncertainty requires exponentially greater computation.”³⁸

As radical uncertainty increases, the digital transformation is meanwhile expanding the domain of human judgment and trust. Institutional economics has generally considered two modes of organization: the market, in which price allocates resources, and the hierarchy, in which authority and contract apply. But neither price nor authority function well as allocation mechanisms when knowledge-based assets are important in production.³⁹ That the market is a poor vehicle for the use and provision of public goods such as knowledge is a standard piece of economic theory. Similarly, a large body of management literature notes that knowledge is hoarded at the top of hierarchical organizations, which are consequently good at routine activities but not at adaptation or innovation.

Trust is a more effective mechanism than either market exchange or command-and-control for coordinating knowledge-intensive activities, both within organizations and between them. The economics literature has long recognized the challenge of asymmetric information and tacit knowledge.⁴⁰ In the digital knowledge economy, tacit or hard-to-codify knowledge is increasingly important. For example, the advantage of high productivity firms over others is encapsulated in the concept of their “organizational capital.” It reflects their ability to manage a complex and uncertain environment, make use of data and software, and employ skilled people who have the authority to make decisions. The gap between firms with high organizational capital and others is growing.⁴¹ Trust networks or communities need to join market and hierarchy as a standard organizational form. Trust is also essential when questions of accountability are blurred, as is the case with hard-to-audit automated-decision systems; the alternative is costly insurance and/or litigation to assign responsibility for outcomes.

The desire for the seeing room view rests on an assumption about the possibility of classifying the world and ordering data as statistical inputs for that synoptic view. Big data does not help overcome the limitations of having to impose a classification: AI techniques involve the aggregation of the vast quantities of raw, irregular, often by-product data into lower dimensional constructs. The machine is doing the classification in ways not legible to humans, but it is doing so nonetheless. But there is much useful knowledge that is tacit rather than explicit and therefore impossible to classify. There is much that is highly locally heterogeneous such that population averages mislead. Nor does having big data and AI overcome the inevitable clash of values or interests that arise in any specific decision-making context. Algorithms cannot adjudicate trade-offs and conflicts; only humans can do so with any legitimacy.

We should think of machines and humans as complements. As societal complexity and uncertainty increase, and as the zone of automated decisions expands, this requires more use of human judgment, not less. Otherwise, we will end up with Scott's disasters of modernism, fully automated. Practical, tacit, improvisational knowledge and informal decision-making processes are always essential for actions to deliver better outcomes locally: even setting aside the point that people might have different and irreconcilable views about what constitutes "better," there are limits to classifiable knowledge, and limits to data.

The use of AI in society must reflect the social nature of data. Although big data offers great potential for progress, any data set is a limited, encoded representation of reality, embedding biases and assumptions, and ignoring information that cannot be codified. A synoptic view of society from a data-enabled seeing room is impossible because no authority can stand outside the reality their decisions will in fact shape. For the promise of AI to be realized, three things are needed: new norms (as well as laws and technologies) governing access and use of data, embedding offline limits online; effective organizations empowering human judgment alongside automated decisions; and legitimate processes to shape the collective decisions being coded into AI. Adopting AI first and reflecting on these needs later is the wrong way to go about socializing data.

AUTHOR'S NOTE

My thanks to the following colleagues for their helpful comments on an early draft: Vasco Carvalho, Verity Harding, Bill Janeway, Michael Kenny, Neil Lawrence, and Claire Melamed. I am entirely responsible for any errors or infelicities. Thanks also to Annabel Manley for research assistance.

ABOUT THE AUTHOR

Diane Coyle is the Bennett Professor of Public Policy at the University of Cambridge. She is the author of *Cogs and Monsters: What Economics Is, and What It Should Be* (2021), *Markets, State, and People: Economics for Public Policy* (2020), *GDP: A Brief but Affectionate History* (2014), and *The Economics of Enough* (2011).

ENDNOTES

- ¹ Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, N.J.: Princeton University Press, 1995).

- ² Andrew Whitby, *The Sum of the People: How the Census Has Shaped Nations, from the Ancient World to the Modern Age* (New York: Hachette, 2020).
- ³ James C. Scott, *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed* (New Haven, Conn.: Yale University Press, 2020), 79.
- ⁴ André Vanoli, *A History of National Accounting* (Amsterdam: IOS Press, 2005), 158.
- ⁵ Porter, *Trust in Numbers*, 43.
- ⁶ Diane Coyle, *GDP: A Brief but Affectionate History* (Princeton, N.J.: Princeton University Press, 2014).
- ⁷ Philipp Lepenies, *The Power of a Single Number: A Political History of GDP* (New York: Columbia University Press, 2016); and “Get a Discount with the Eat Out to Help Out Scheme,” Gov.uk, July 15, 2020, updated September 1, 2020, <https://www.gov.uk/guidance/get-a-discount-with-the-eat-out-to-help-out-scheme> (accessed October 29, 2020).
- ⁸ Alain Desrosières, *The Politics of Large Numbers: A History of Statistical Reasoning* (Cambridge, Mass.: Harvard University Press, 1998), 252.
- ⁹ Diane Coyle, *The Weightless World* (Cambridge, Mass.: MIT Press, 1997).
- ¹⁰ Jill Lepore, *If Then: How One Data Company Invented the Future* (New York: Hachette, 2020).
- ¹¹ Eden Medina, *Cybernetic Revolutionaries: Technology and Politics in Allende’s Chile* (Cambridge, Mass.: MIT Press, 2014).
- ¹² Aaron Bastani, *Fully Automated Luxury Communism* (New York: Verso Books, 2019).
- ¹³ “Covid-19 Crisis Accelerates UK Military’s Push into Virtual War Gaming,” *Financial Times*, August 19, 2020, <https://www.ft.com/content/ab767ccf-650e-4afb-9f72-2cc84efao708>.
- ¹⁴ Diane Coyle and Stephanie Diepeveen, “Creating and Governing Value from Data” (2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3973034.
- ¹⁵ Shannon Mattern, “The Spectacle of Data: A Century of Fairs, Fiches, and Fantasies,” *Theory, Culture & Society* 37 (7–8) (2020): 133–155.
- ¹⁶ Dominic Cummings, “On the Referendum #33: High Performance Government, ‘Cognitive Technologies,’ Michael Nielsen, Bret Victor, and ‘Seeing Rooms,’” Dominic Cummings’s Blog, June 26, 2019, <https://dominiccummings.com/2019/06/26/on-the-referendum-33-high-performance-government-cognitive-technologies-michael-nielsen-bret-victor-seeing-rooms/>.
- ¹⁷ Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (London: Profile Books, 2019).
- ¹⁸ Organisation for Economic Co-operation and Development, “Statement on a Two-Pillar Solution to Address the Tax Challenges Arising from the Digitalisation of the Economy,” OECD/G20 Base Erosion and Profit Shifting Project, July 1, 2021, <https://www.oecd.org/tax/beps/statement-on-a-two-pillar-solution-to-address-the-tax-challenges-arising-from-the-digitalisation-of-the-economy-july-2021.pdf>.
- ¹⁹ Darin Bartholomew, “Long Comment Regarding a Proposed Exemption under 17 U.S.C. 1201,” https://copyright.gov/1201/2015/comments-032715/class%2021/John_Deere_Class21_1201_2014.pdf.

- ²⁰ Imanol Arrieta-Ibarra, Leonard Goff, Diego Jiménez-Hernández, et al., “Should We Treat Data as Labor? Moving beyond ‘Free,’” *AEA Papers and Proceedings* 108 (2018): 38–42; and Eric A. Posner and E. Glen Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society* (Princeton, N.J.: Princeton University Press, 2019).
- ²¹ Zoë Hitzig, Lily Hu, and Salomé Viljoen, “The Technological Politics of Mechanism Design,” *University of Chicago Law Review* 87 (1) (2019), <https://ssrn.com/abstract=3638585>.
- ²² Diane Coyle, Stephanie Diepeveen, Julia Wdowin, et al., *The Value of Data: Policy Implications – Main Report* (Cambridge: Bennett Institute for Public Policy Research, 2020), <https://www.bennettinstitute.cam.ac.uk/publications/value-data-policy-implications/>.
- ²³ Salomé Viljoen, “Democratic Data: A Relational Theory for Data Governance,” *Yale Law Journal* 131 (2020), <https://ssrn.com/abstract=3727562>.
- ²⁴ Herbert A. Simon, “A Behavioral Model of Rational Choice,” *The Quarterly Journal of Economics* 69 (1) (1955): 99–118.
- ²⁵ Robert E. Lucas, “Econometric Policy Evaluation: A Critique,” *Carnegie-Rochester Conference Series on Public Policy* 1 (1) (1976): 19–46.
- ²⁶ Ronald M. Lee, “Bureaucracies, Bureaucrats and Information Technology,” *European Journal of Operational Research* 18 (1984): 293–303.
- ²⁷ Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, et al., “A Survey on Bias and Fairness in Machine Learning,” arXiv (2019), <https://arxiv.org/abs/1908.09635v2>; and Xavier Ferrer, Tom van Nuenen, José M. Such, et al., “Bias and Discrimination in AI: A Cross-Disciplinary Perspective,” arXiv (2020), <https://arxiv.org/abs/2008.07309>.
- ²⁸ Cass R. Sunstein, “Incompletely Theorized Agreements,” *Harvard Law Review* 108 (7) (1995): 1733–1772.
- ²⁹ Diane Coyle and Adrian Weller, “‘Explaining’ Machine Learning Reveals Policy Challenges,” *Science* 368 (6498) (2020): 1433–1434.
- ³⁰ See, for example, Thomas Arnold, Daniel Kasenberg, and Matthias Scheutz, “Value Alignment or Misalignment—What Will Keep Systems Accountable?” in *AI, Ethics, and Society: Papers from the 2017 AAAI Workshop, San Francisco, California, USA, February 4, 2017*, ed. Toby Walsh (Menlo Park, Calif.: AAAI Press, 2017); and Iason Gabriel, “Artificial Intelligence, Values, and Alignment,” *Minds and Machines* 30 (2020): 411–437, <https://doi.org/10.1007/s11023-020-09539-2>.
- ³¹ Stuart Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (New York: Penguin, 2019).
- ³² Gabriel, “Artificial Intelligence, Values, and Alignment.”
- ³³ Coyle and Diepeveen, “Creating and Governing Value from Data.”
- ³⁴ Georg Simmel, “The Sociology of Secrecy and of Secret Societies,” *American Journal of Sociology* 11 (4) (1906): 441–498.
- ³⁵ Richard Warner and Robert H. Sloan, “The Self, the Stasi, and NSA: Privacy, Knowledge, and Complicity in the Surveillance State,” *Minnesota Journal of Law, Science & Technology* 17 (2016): 347.
- ³⁶ Linnet Taylor, “The Ethics of Big Data as a Public Good: Which Public? Whose Good?” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083) (2016).

- ³⁷ John Kay and Mervyn King, *Radical Uncertainty: Decision-Making Beyond the Numbers* (New York: W. W. Norton & Company, 2020).
- ³⁸ Neil Lawrence, “Future of AI 6. Discussion of ‘Superintelligence: Paths, Dangers, Strategies,’” *Inverseprobability* blog, May 9, 2016, <https://inverseprobability.com/2016/05/09/machine-learning-futures-6>.
- ³⁹ Paul S. Adler, “Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism,” *Organization Science* 12 (2) (2001): 215–234, <https://doi.org/10.1287/orsc.12.2.215.10117>.
- ⁴⁰ See, for example, Sanford Grossman and Joseph E. Stiglitz, “Information and Competitive Price Systems,” *The American Economic Review* 66 (2) (1976): 246–253; Bengt Holmstrom, “The Firm as a Subeconomy,” *Journal of Law, Economics and Organization* 15 (1) (1999): 74–102, <https://doi.org/10.1093/jleo/15.1.74>; and Luis Garicano and Esteban Rossi-Hansberg, “Organization and Inequality in a Knowledge Economy,” *The Quarterly Journal of Economics* 121 (4) (2006): 1383–1435.
- ⁴¹ Lorin M. Hitt, Shinkyu Yang, and Erik Brynjolfsson, “Intangible Assets: Computers and Organizational Capital,” *Brookings Papers on Economic Activity* 1 (2002): 137–181; and Prasanna Tambe, Lorin Hitt, Daniel Rock, and Erik Brynjolfsson, “Digital Capital and Superstar Firms,” NBER Working Paper No. 28285 (Cambridge, Mass.: National Bureau of Economic Research, 2020).