MATTHEW L. JONES*

How We Became Instrumentalists (Again): Data Positivism since World War II

ABSTRACT

In the last two decades, a highly instrumentalist form of statistical and machine learning has achieved an extraordinary success as the computational heart of the phenomenon glossed as "predictive analytics," "data mining," or "data science." This instrumentalist culture of prediction emerged from subfields within applied statistics, artificial intelligence, and database management. This essay looks at representative developments within computational statistics and pattern recognition from the 1950s onward, in the United States and beyond, central to the explosion of algorithms, techniques, and epistemic values that ultimately came together in the data sciences of today. This essay is part of a special issue entitled *Histories of Data and the Database* edited by Soraya de Chadarevian and Theodore M. Porter.

KEY WORDS: computational statistics, data mining, data sciences, big data, pattern recognition, instrumentalism, Leo Breiman

In the last two decades, a highly instrumentalist form of statistical and machine learning has achieved an extraordinary success as the computational heart of the phenomenon glossed as "predictive analytics," "data mining," or "data science." The current movement of data-focused computational analysis has emerged from the loose confederating of a range of areas of inquiry focused on data that developed through the Cold War on both sides of the Iron Curtain, domains that have exploded in the commercial, national security, and academic worlds since the early 1990s. Over the years, investigators working in these areas of research have made heavy use of mathematical statistics, while breaking from the values, training, procedures, publication patterns, and

*Department of History, Columbia University, 514 Fayerweather Hall—MC2513, 1180 Amsterdam Ave., New York, NY 10027, mj340@columbia.edu

Historical Studies in the Natural Sciences, Vol. 48, Number 5, pps. 673–684. ISSN 1939-1811, electronic ISSN 1939-182X. © 2018 by the Regents of the University of California. All rights reserved. Please direct all requests for permission to photocopy or reproduce article content through the University of California Press's Reprints and Permissions web page, http://www.ucpress.edu/journals.php?p=reprints. DOI: https://doi.org/10.1525/hsns.2018.48.5.673.

funding of that academic field. They have equally sought to adapt algorithms for working with complex and increasingly large real-world data sets while reducing computational costs and delays.

Researchers in many of these areas have long prized prediction over knowledge about an interpretable, causal, or mechanistic model.[1] In 2001, the renegade statistician Leo Breiman described "two cultures" of using statistical models "to reach conclusions from data." The first "assumes that the data are generated by a given stochastic data model." The second "uses algorithmic models and treats the data mechanisms as unknown."[2] This essay sketches the diverse sources of this second algorithmic culture, one stemming more from an engineering culture of predictive utility than from a scientific culture of truth.

From the late 1940s onward, much of this work happened at the peripheries of well-established academic disciplines, from electrical engineering to physics, and much of it within research units heavily supported by military and intelligence funding. Decades before they allowed Google to search and Amazon to recommend, highly instrumentalist predictive algorithms running on robust computational platforms learned to read the numbers printed on checks and worked to discern planes and tanks in reconnaissance photos.

This instrumentalist culture of prediction has emerged from subfields within applied statistics, artificial intelligence, and database management.[3] In this short essay, I focus on the first, statistics, and look at a few representative developments in computational statistics central to the explosion of algorithms—techniques and epistemic values that ultimately came together in the data sciences of today. This essay focuses on topics and actors quite distinct from the better known and, until recently, higher status fields of artificial intelligence, expert systems, and operations research.[4] The paper

1. For recent studies of cultures of prediction, see Ann Johnson, "Rational and Empirical Cultures of Prediction," in *Mathematics as a Tool*, eds. Johannes Lenhard and Martin Carrier, vol. 327 (Cham: Springer International Publishing, 2017), 23–35, https://doi.org/10.1007/978-3-319-54469-4_2. Matthias Heymann, Gabriele Gramelsberger, and Martin Mahony, *Cultures of Prediction in Atmospheric and Climate Science: Epistemic and Cultural Shifts in Computer-Based Modelling and Simulation* (London: Routledge, 2017).

2. Leo Breiman, "Statistical Modeling: The Two Cultures," *Statistical Science* 16, no. 3 (2001): 199.

3. Excluded here likewise is financial prediction, among others; for starting points, see Walter A. Friedman, *Fortune Tellers: The Story of America's First Economic Forecasters* (Princeton, NJ: Princeton University Press, 2014).

4. For the sparsity of histories of machine learning, see Aaron Plasek, "On the Cruelty of Really Writing a History of Machine Learning," *IEEE Annals of the History of Computing* 38, no. 4

considers first the growth, in the 1950s and 1960s, of two strands of computational statistical approaches, exploratory data analysis and pattern recognition, that focused on the increasing explosion of data during the Cold War. The essay then discusses parallel movements outside of the United States, before looking at the intensifying consolidation of an instrumentalist statistical approach as central to contending with ever-expanding data sets, commercial and governmental, from the 1990s to the present.

MATHEMATICAL STATISTICS AND ITS DISCONTENTS

In a 1962 manifesto, the Princeton-Bell Labs mathematician John Tukey called for new approach he dubbed "data analysis" that would be dedicated as much to discovery as to confirmation:

> Data analysis, and the parts of statistics which adhere to it, must then take on the characteristics of a science rather than those of mathematics, specifically:
>
> (1) Data analysis must seek for scope and usefulness rather than security.
> (2) Data analysis must be willing to err moderately often in order that inadequate evidence shall more often suggest the right answer.
> (3) Data analysis must use mathematical argument and mathematical results as bases for judgment rather than as bases for proofs or stamps of validity.[5]

As a scientific practice, Tukey continued, data analysis is an art, not a logically closed discipline. Tukey was crystallizing an alternate approach to academic statistics, one that used the mathematical power of statistical thinking for exploratory as much as confirmatory purposes, and one that might be applicable to observed data, not exclusively to experimental data produced as part of an experimental trial. Thanks to "war problems" in the 1940s, Tukey explained in an interview, "it was natural to regard statistics as something that had the purpose of being used on data—maybe not directly, but at most at some remove. Now, I can't believe that other people who had practical experience failed to have this view, but they certainly—I would say—failed to

---

(Dec 2016): 6–8, https://doi.org/10.1109/MAHC.2016.43. For new histories of artificial intelligence, see Stephanie Dick, "Of Models and Machines: Implementing Bounded Rationality," *Isis* 106, no. 3 (2015): 623–34.

5. John W. Tukey, "The Future of Data Analysis," *Annals of Mathematical Statistics*, no. 1 (Mar 1962): 1–67, on 6; https://doi.org/10.1214/aoms/1177704711.

advertise it."[6] Indeed, thanks to the support of Mina Rees and the efforts of the statistician Harold Hotelling and others, the great successes of highly *applied* statistics during WWII were channeled into financial and symbolic support for the creation of a mathematically focused, theoretical statistics in the United States and in Europe, rather than a more practically oriented, data-focused statistics.[7] Before long, in the eyes of critics such as Tukey, practical data collection and analysis were sacrificed at the altar of mathematical sophistication and rigor. Tukey and other critics complained that relatively few within academic mathematical statistics and its allied branches, such as econometrics, celebrated the practical cultivation of data analysis and forms of judgment as a central endeavor. Data analysis flourished elsewhere, in the penumbra of mathematical statistics and other well-established disciplines, in corporate research labs and engineering departments, under various names.

PATTERN RECOGNITION

In the early 1960s, engineers at Philco, newly a division of Ford Motors, worked under contract with the U.S. Army on technological means to aid the military in the automated recognition of features in photos. Among the bevy of technologies the Army supported, at least one group was using computational statistics to aid classification.[8] In just such commercial and academic labs funded by the U.S. military and intelligence agencies did forms of computational statistics focused more on predictions based on data and less on the confirmation of causal hypotheses flourish. Researchers such as the Philco engineers working within the broad rubric of "pattern recognition" sought techniques to discriminate among objects, estimating parameters for known distributions, and hardest of all, to begin the tough task of discerning

6. Luisa T. Fernholz et al., "A Conversation with John W. Tukey and Elizabeth Tukey," *Statistical Science* (2000): 80–81.

7. Patti W. Hunter, "Drawing the Boundaries: Mathematical Statistics in 20th-Century America," *Historia Mathematica* 23, no. 1 (1996): 7–30. For nuanced treatments of wartime developments see, e.g., Judy L. Klein, "Economics for a Client: The Case of Statistical Quality Control and Sequential Analysis," *History of Political Economy* 32, Suppl. no. 1 (2000): 25–70. More broadly, see T. Dryer, "Algorithms under the Reign of Probability," *IEEE Annals of the History of Computing* 40, no. 1 (Jan 2018): 93–96, https://doi.org/10.1109/MAHC.2018.012171275.

8. T. Harley, J. Bryan, and L. Kanal, "Semi-Automatic Imagery Screening Research Study and Experimental Investigation," Philco Advanced Technology Laboratory, Blue Bell, PA, 29 Mar 1963.

probability distributions when their underlying form cannot be assumed.[9] They worked at government labs, at corporate labs, and at universities such as Cornell, U.S.C., and Stanford, typically with copious military support.[10] Although these efforts began in some cases with special purpose physical machines to perform classification, researchers increasingly made digital computers applying statistical methods more central. When they surveyed the field in the 1960s and early 1970s, researchers explained that pattern recognition involved less an academic discipline than a cluster of like-minded practitioners oriented around common sets of goals. The neural network idea of the perceptron is perhaps the best known of these efforts. By the late 1960s, most researchers in pattern recognition ultimately cared little whether neural networks in any way replicated human cognition; the networks were tools for prediction, not means for understanding the brain: "Whatever success we have had [has] been the result of an effective transformation of a perception-recognition problem into a classification problem."[11]

In the course of this work, early forms of many of the key algorithms now central to the contemporary data sciences emerged; researchers modified these algorithms to work within the computational limits of their times on dirty, real-world data sets. "Practical considerations of computer economics often prevent the wholesale application of the methods mentioned above to real-life situations." Such situations require "somewhat undignified and haphazard manipulation . . . to render the problem amenable to orderly solution," including "preprocessing, filtering or prefiltering, feature or measurement extraction, or dimensionality reduction."[12] Techniques for handling troublesome data in existing hardware were integral, not ancillary, to pattern recognition in practice.

The need for large-scale data storage became apparent soon after the Second World War within the American intelligence community. While supporting IBM's development of larger data storage, the National Security Agency organized key early conferences to encourage the development of robust database

9. Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge; New York: Cambridge University Press, 2010), ch. 4.

10. For the importance of this gray area of research between government and academia, see Joy Rohde, *Armed with Expertise: The Militarization of American Social Research during the Cold War* (Ithaca, NY: Cornell University Press, 2013).

11. Laveen N. Kanal, "Preface," in *Pattern Recognition*, ed. Laveen N. Kanal (Washington, DC: Thompson Book Co., 1968), xi.

12. G. Nagy, "State of the Art in Pattern Recognition," *Proceedings of the IEEE* 56, no. 5 (May 1968): 836, https://doi.org/10.1109/PROC.1968.6414.

solutions. Starting with the data from airplane reservation systems in the 1960s, industry began accumulating data about customers at a rapidly accelerating rate. Much of the interesting data is "big" in two different ways: it involves observations about, say, a large number of people or a large number of purchases; and it involves, for each one of those observations, a large number of variables. The last point—called the "high-dimensionality" of data—involved major mathematical and computational work, far from trivial, but as often distant from the concerns of academic statisticians.

DATA-FOCUSED STATISTICS OUTSIDE THE UNITED STATES

Data-driven computational statistics developed in tension with mathematical statistics outside the Anglophone world. In France, J.-P. Benzécri created a powerful school of "analyse des données" focused on more powerful exploratory and descriptive statistics using computers. "The progress of the 'analyse des données' due to computers," he wrote, "will not continue without upsetting all of statistics."[13] In Japan, Hayashi Chikiō developed a set of practices he named *Deta no Kagaku*, the "Science of Data," as an alternative to mathematical statistics, which he described as "good-for-nothing and not understandable."[14]

Developments in the Soviet Union were probably of the most consequence to the recent history of machine learning and computational statistics. In 2006, the machine-learning specialist Vladimir Vapnik explained the revolutionary transformation of computer learning in the Soviet Union. In the statistical approach dominant in the middle of the twentieth century, he detailed, a "generative model of induction" predominated, where "an understanding of how data are generated reflects an understanding of the corresponding law of nature." In data sets with high numbers of dimensions, such approaches turned out to fail. In their place Vapnik and like-minded colleagues created "the predictive (discriminative) models of induction." In such an approach, "they are just looking for a function that explains the data best."[15]

13. See Alain Desrosières, *Prouver et gouverner: une analyse politique des statistiques publiques* (Paris: Découverte, 2014), ch. 9.

14. Hayashi Chikiō, in C. Hayashi and M. Takahashi, "Kagakusi to Kagakusha: Hayashi Chikiosi Kōkai Intabyu," *Kōdō Keiryōgaku* 31, no. 2 (2004): 107–24, quoted and translated in Joonwoo Son, "Data Science in Japan" (unpublished ms., May 2016), 1.

15. Vladimir Naumovich Vapnik, *Estimation of Dependences Based on Empirical Data (1982); Empirical Inference Science: Afterword of 2006*, 2nd ed. (New York: Springer, 2006), 415.

In terms of the philosophy of science, he explained, the generative approach is realist, and the predictive one, instrumentalist. This instrumental approach, moreover, "played a crucial role in the success that pattern recognition technology has achieved."[16]

Although he spent much of his later career in the United States at Bell Labs, Vapnik came to this instrumentalist approach, and the high-dimensional data sets, as a member of the Institute for Control Science of the Academy of Sciences of the U.S.S.R. in the 1960s and 1970s.[17] The Institute allowed the flourishing of a highly computationally focused learning approach using data of large size and dimensionality. In the U.S.S.R. as in the United States, pattern recognition researchers understood themselves as distant from the ambitions and rigors of symbolic artificial intelligence and classical academic statistics, even as they drew heavily upon practices and ideas from both.[18]

The U.S. and Soviet strands came together at Bell Labs in the 1990s. The technique most associated with Vapnik, Support Vector Machines, came to fruition in a remarkable collaboration there, where he joined forces with the French neural net researcher Isabella Guyon, among many others. Like other major examples of development in the computational data sciences, Vapnik worked under the imperative of contending with high-dimensional data within a funding regime supporting it, and without the necessity of producing symbolic artificial intelligence or causal models.[19]

BEYOND PATTERN RECOGNITION: BREIMAN AND THE DEVELOPMENT OF DECISION TREES

In the 1970s, Leo Breiman was trying to predict Los Angeles pollution patterns for the Environmental Protection Agency and to discern ship profiles for what he termed "spook agencies" using large amounts of high-dimensional data. In

16. Ibid., 417.

17. See "History of the Institute | ИПУ РАН," http://www.ipu.ru/en/node/12744 (accessed 7 Jul 2017).

18. See the discussion in Ya.Z. Tsypkin, ed., "Adaptation and Learning," 44–75, on 4, ch. 3 in *Adaptation and Learning in Automatic Systems*, vol. 73 of *Mathematics in Science and Engineering* (Elsevier, 1971), https://doi.org/10.1016/S0076-5392(08)62696-X.

19. On the importance of this Soviet work, see Léon Bottou, "In Hindsight: Doklady Akademii Nauk SSSR, 181 (4), 1968," in *Empirical Inference*, eds. B. Schölkopf, Z. Luo, and V. Vovk (Berlin, Heidelberg: Springer, 2013), 3–5, http://link.springer.com/chapter/10.1007/978-3-642-41136-6_1.

the 2000s, he published his two-cultures manifesto calling for academic statisticians to drop their insistence on data models and develop tools more suited to the complexities of real-world data. In the interim, he had become a key progenitor of a family of algorithms called decision trees and a development of them he christened "random forests" that as of 2018 is among the most predictive and, it appears, widely used of all machine-learning algorithms.

In his consulting work, Breiman found himself wanting to draw upon the complexity of his data sets without having to reduce their dimensionality. "In the usual pattern recognition approach," he and a colleague explained, the usual algorithms "give a one gulp answer" that requires a "drastic reduction in dimensionality . . . to make the sample size sufficiently dense in the space to define the problem and to make it computationally feasible."[20] Reduction was a one-way street: "The loss in information is irrevocable."[21]

In tandem with several other researchers working around the same moment, Breiman came to celebrate decision trees as superior to existing pattern recognition techniques. The excitement of trees for him was that they dealt with large amounts of data piecemeal, not all at once.[22] The algorithm leveraged, rather than reduced, the high-dimensional real-world data.

DATA MINING IN THE 1990s: PATTERN RECOGNITION APPLIED AND REBRANDED

By the late 1980s, the tools for analyzing this stored business data came to be seen as increasingly inadequate. Similar stories held true with scientific, military, and intelligence data. In 1998, amid the blossoming of large-scale corporate, government, and academic "data warehouses," a key researcher explained, "A large data store today, in practice, is not very far from being a grand, write-only, data tomb."[23] Much needed to be done.

The movement known as "data mining" emerged in the early 1990s to make sense of the rapidly growing untapped stores of corporate and scientific

20. William S. Meisel and Leo Breiman, "Topics in the Analysis and Optimization of Complex Systems. Appendix B. Tree Structured Classification Methods," Final report to AFOSR (Technology Service Corporation, 28 Feb 1977), 4, http://www.dtic.mil/dtic/tr/fulltext/u2/a038209.pdf.

21. Ibid.

22. Ibid.

23. Usama Fayyad, "Mining Databases: Towards Algorithms for Knowledge Discovery," *Bulletin of the Technical Committee on Data Engineering* 21, no. 1 (1998): 48.

data using the tools produced in pattern recognition research, other artificial intelligence approaches, and applied computational statistical work, such as the work of Breiman and Tukey. Data mining, or more formally, Knowledge Discovery in Databases (KDD), is the activity of creating non-trivial knowledge suitable for action from databases of vast size and dimensionality.[24] Data mining concerns databases of very large size—millions or billions of records, usually with elements of high dimensionality. Performing reasonably fast analyses of high-dimensional, messy, real-world data is central to the identity and purpose of data mining, even more so than in pattern recognition. Sophisticated statistical and machine learning algorithms were typically devised for sets of data that can easily fit in memory, or that require a relatively small use of slower disk access. Adapting such algorithms to huge quantities of data that cannot be held in memory is non-trivial.[25] With some key exceptions, such as Breiman, theoretically oriented practitioners did not focus upon this computational challenge; other business and scientific communities made it a central concern.[26]

STATISTICAL MODERNIZATION

In the late 1970s, a small but steadily increasing number of computationally minded statisticians called for their field to more fully embrace the possibilities the digital computer afforded, in graduate training as well in the understanding of what comprised good science. In 1979, Stanford statistician Bradley Efron called for his colleagues to recognize that computers had dramatically transformed what constituted "simple" in a theory. The computer, he argued, "has redefined 'simple' in the mathematical sciences." Mathematics correspondingly had to shift. The avalanche of data would require "a blend of traditional mathematical thinking combined with the numerical and organizational

24. Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, "From Data Mining to Knowledge Discovery: An Overview," in *Advances in Knowledge Discovery and Data Mining* (Menlo Park, CA: AAAI/MIT Press, 1996), 1–34.

25. Compare computation and data friction as discussed in Paul Edwards, *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming* (Cambridge, MA: MIT Press, 2010).

26. Matthew L. Jones, "Querying the Archive: Database Mining from Apriori to Page-Rank," in *Science in the Archives: Pasts, Presents, Futures*, ed. Lorraine Daston (Chicago: Chicago University Press, 2016), 311–28.

aptitude of the computer."[27] Soviet scholars had made much the same point in the early 1970s.

Practitioners such as Breiman, Efron, and Bell Labs' William Cleveland argued that academic statisticians had failed to face up to large real-world data sets and to integrate computing more centrally within their understanding of the field. In a 1993 call for a "greater statistics" that would learn from data, John Chambers of Bell Labs worried that the overly insular mathematical drive of statistics was "limiting both the influence of statistics and the benefits the field had provided to society."[28] In 2002, the U.S. National Science Foundation funded a workshop on the future of statistics that led to calls for a dramatic transformation of the discipline to retain its centrality amid the explosion of data and the threat of competitor disciplines focused on analyzing data.[29] The explosion of data had created the opportunity for statistics to serve an essential regulatory function, but the field as a whole was failing to seize the opportunity. Responding to the report, two major computationally oriented statisticians, Walter Stuetzel and David Madigan, called for an upending of graduate education in their field: "Statistics as a discipline exists to develop tools for analyzing data. As such, statistics is an engineering discipline and methodology." Curricula remained old-fashioned, too mathematical.

> Statistics has primarily focused on squeezing the maximum amount of information out of limited data. This paradigm is rapidly diminishing in importance and statistics education finds itself out of step with reality.[30]

Breiman was more blunt: the NSF report "denigrates the way that the most important advances in statistics have occurred—not by introspection, but involvement in challenging problems suggested by different disciplines." The report, he explained, "is a step into the past . . . ," back into a mathematically solipsistic dark age.[31]

27. Bradley Efron, "Computers and the Theory of Statistics: Thinking the Unthinkable," *SIAM Review* 21, no. 4 (1979): 480.

28. John M. Chambers, "Greater or Lesser Statistics: A Choice for Future Research," *Statistics and Computing* 3, no. 4 (1993): 182.

29. Bruce G. Lindsay, Jon Kettenring, and David O. Siegmund, "Statistics: Challenges and Opportunities for the Twenty-First Century," 20 Jun 2003, https://web.archive.org/web/20040707164725/http://www.stat.psu.edu:80/bgl/nsf_report.pdf.

30. David Madigan and Werner Stuetzle, "[A Report on the Future of Statistics]: Comment," *Statistical Science* 19, no. 3 (2004): 408.

31. Leo Breiman, "[A Report on the Future of Statistics]: Comment," *Statistical Science* 19, no. 3 (2004): 411.

ENSEMBLE MODELING AND THE PREDICTIVE ETHOS

By the late 1990s, a growing literature began to show, Leo Breiman argued, that "combining a multiple set of predictors, all constructed using the same data, can lead to dramatic decreases in test error." This predictive success came at great cost. "At the end of the day, what we are left with is an almost inscrutable prediction function combining many different predictors. But the resulting predictor can be quite accurate."[32] Despite their epistemologically questionable status, such inscrutable combinations predict better. A bevy of techniques with snappy names emerged to create such ensembles: bagging, boosting, arcing, etc.

The predictive gains were massive, and to an increasing number of practitioners among different scientific, intelligence, and business domains, these gains have come to overshadow the massive opacity of the predictive ensemble generated. Increasingly, practitioners have abandoned using any one family of predictive models in favor of combining many different predictors, most famously exemplified in the victor of the Netflix Prize in 2009. This dramatic success of ensembles amplified the ethic of prediction over interpretation. A generation ago, the inscrutability of neural nets made them deeply problematic; the renaissance of neural networks from around 2012 rests squarely on the legitimation of such ensemble models, for commerce, for spooks, and for science.[33]

SCIENCES OF THE PARTICULAR

For at least three decades, professional historians of science have pushed against a vision of science modeled on theoretical physics; we now celebrate the diverse forms of knowledge focused on the careful empirical study of particular things. Exponents of data-focused computational science have a surprisingly similar evolution. Just as history of science embraced the study of the particular as it disconnected from a Cold War prioritizing of theory, the data sciences moved beyond the aggregates of mathematical statistics to draw upon granular data sets to characterize particular things—individual people, diseases, films. In a key manifesto celebrating the "unreasonable effectiveness

32. Leo Breiman and Nong Shang, "Born Again Trees," University of California, Berkeley, Berkeley, CA, Technical Report, 1996.

33. For resistance to this instrumentalist focus, see the remarkable Judea Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge: Cambridge University Press, 2000).

of data," three Google researchers argued in terms that echo humanist denunciations of reductionist knowledge: "sciences that involve human beings rather than elementary particles have proven more resistant to elegant mathematics." Something else is needed. "Perhaps when it comes to natural language processing and related fields, we're doomed to complex theories that will never have the elegance of physics equations. But if that's so, we should . . . embrace complexity and make use of the best ally we have: the unreasonable effectiveness of data."[34]

Much of the promise of the data sciences, whether in medicine, marketing, or getting out the vote, ostensibly comes from overcoming older theory-laden categorizations to characterize individuals in their specificity, all in order to predict their behavior. This positivism has, unsurprisingly, resulted in new algorithmic processes strengthening traditional categories of racial, sexual, and class discrimination.[35] The long-fought humanist desire to focus on the individual has, in a most peculiar turnabout, encountered perhaps the most powerful system of manipulating human emotions at scale the world has yet seen.

ACKNOWLEDGEMENTS

Thanks to Stephanie Dick, Tom Stapleford, Joy Rohde, Michael Gordin, Chris Wiggins, Ted Porter, and Soraya de Chadarevian.

34. A. Halevy, P. Norvig, and F. Pereira, "The Unreasonable Effectiveness of Data," *Intelligent Systems, IEEE* 24, no. 2 (April 2009): 8–12, https://doi.org/10.1109/MIS.2009.36.

35. See recently Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).