Matthew L. Jones

AI in History

Late in 1982, Edinburgh professor Donald Michie explained the fundamental error that plagued earlier efforts to create artificial intelligence. "The inductive learning of concepts, rules, strategies, etc. from *examples* is what confers on the human problem-solver his power and versatility, and not (as had earlier been supposed) power of calculation."[1] A minority position in 1982, learning from examples came to dominate artificial intelligence early in the new millennium. In a key 2009 manifesto celebrating the "unreasonable effectiveness of data," three Google researchers argued "sciences that involve human beings rather than elementary particles have proven more resistant to elegant mathematics." We should "embrace complexity and make use of the best ally we have: the unreasonable effectiveness of data."[2]

The computer scientist John McCarthy coined the term "artificial intelligence" originally in search of funding; in the mid 2010s, the term was dramatically redefined to describe large-scale algorithmic decision-making systems and predictive machine learning "trained" on massive data sets.[3] Through most of the Cold War and beyond, AI researchers focused on "symbolic AI" largely ignored data collected from everyday and military activities.[4] Such empiricism of the quotidian paled in prestige in comparison with logic and numerical computation and the more empirically oriented approaches such as neural networks and pattern recognition were widely lambasted.[5] Learning from data seemed to be the wrong approach for producing intelligence or intelligent behaviors. Alongside this symbolic approach, in the USA, USSR, and beyond, a far less prestigious empiricist stratum developed comprising congeries of techniques for dealing with large-scale military, intelligence, and commercial data. Our contemporary world of AI, with its often-biased algorithmic decision system, owes far more to this empirical strand of inquiry than to the previously higher status and much studied symbolic artificial intelligence.[6]

Data-focused machine learning, the artificial intelligence dominant today, takes on board the inability of earlier research to contend with the richness of human experience and the built and natural world. Whether in medicine or advertising, the promise of these new forms of artificial intelligence grounded in everyday data stems from understanding people less as variations within traditional racial or social groups but more precisely as fitting within more nuanced, more empirically driven classifications, with the potential to grasp people, including human language, in their specificity.[7] For all the mathematical and computational complexity involved, machine learning systems instantiate an empiricist temper regnant. Indeed, many of the major problems of these systems—their speed in learning and then reproducing racist speech, for example—comes from their power of memorizing and then predicting the worst as well as the best of human beings and our structures of power and meaning. Oddly enough, machine learning on enormous data sets is a science of particulars, more like, say, history than physics. With its data-driven, rather than reductionist, focus, this form of artificial intelligence threatens precisely the professions centered on particularity, on granular empirical knowledge, clinical knowledge, legal knowledge—and potentially historical knowledge.

How might understanding the history of recent artificial intelligence allow us to draw upon it more reflectively in our historical practice, as a supplement to other digital and traditional practices?

What role do historians have in understanding this newer artificial intelligence and how might we draw upon it? How might our historical understanding of artificial intelligence allow us to use it as historians on our own terms? How might understanding the history of recent artificial intelligence allow us to draw upon it more reflectively in our historical practice, as a supplement to other digital and traditional practices?

Until recently the history of artificial intelligence focused primarily upon its intellectual and technical developments, its relations to cognitive science, and the history of science fiction from Mary Shelley to Philip K. Dick to the Wachowski sisters.[8] Scholars today are excavating the global histories of diverse ways of making machines that emulate facets of human intelligence, involving actors far beyond the usual stories of cybernetics and Alan Turing.[9] Historians, sociologists, and anthropologists have long studied these aspirational efforts to incorporate machines into decision-making, political, educational, and other processes.

While earlier forms of artificial intelligence have had limited impact on social, economic, and political life, the impact of contemporary AI systems is growing rapidly—maybe even exponentially.[10] The automated decision systems driven by contemporary artificial intelligence are ever more woven into the fabric of our interaction with governments, corporations, and each other: they direct and misdirect, they guide and misguide, they support and aid. They are no neutral substratum: they divide and unify, they spy and protect, they target and mistarget. Historians working on the 2000s will not be able to detach political, social, and cultural history from the history of AI systems without losing sight of major causal forces and intermediators of our time.

Grasping the conditions enabling contemporary machine learning systems will require a thicket of diverse histories, produced by historians fruitfully at odds with one another in prioritizing the importance of the social, the economic, the cultural, and the technological. These diverse historical approaches both reveal the structures making data-focused decision systems possible and underscore the imperatives, biases, and interests behind large-scale data collection and analysis. Important recent examples speak to the diversity of approaches needed:

- The limitations of privacy around data since the 1970s.[11]
- Accounts of systemic inequality built into earlier technical and bureaucratic systems.[12]
- Long histories of data's centrality for race, power, and empire.[13]
- Systems of classifications and their exclusions.[14]
- Development of new algorithms and the challenges of scaling them to huge data.[15]
- Infrastructures for recording, long-term transmission, and storage.[16]
- Environmental histories, including water for cooling and inexpensive electricity, and the supply of rare earth minerals.[17]
- Venture capital and tax-avoidance strategies.[18]
- Labor history, including making, editing, and curating data.[19]
- Colonial and neocolonial data extraction and questions of ownership.[20]
- Paradigmatic tales of deskilling—now applied to the professional classes.[21]
- Data visualization and rhetoric around decision-making.[22]

These multiple forms of historicization of AI systems undergird the slow but real transformation underway whereby endeavors long thought the exclusive domain of human beings—the clinical judgment of the radiologist, connoisseurship, or hermeneutics—come increasingly to be computationally tractable. Our vocation as historians—of showing that things are not necessarily so—is vital for understanding just the contingency of the entire array of structures—physical, intellectual, social—upon which artificial intelligence has been built.

Political, cultural, social, and economic historians studying the very late twentieth and twenty-first centuries will increasingly need to weigh the potential effects of large-scale machine learning on politics, society, and science alike—such as the potential causal effects of automatic recommendation systems in fueling the polarization in Myanmar, the UK, and US presidential elections since 2015.[23] Attributing too much about the explosion of right-wing groups to recommendation algorithms misses the rich, long-term histories of white nationalism; yet not to be able to contend with the transformation of scale and new forms of communication, solidarity, and organization machine learning enables will profoundly limit historical inquiry. In the end, AI systems may be much more efficient than final causes of much that we need to explain, but we will need the historical power to investigate mechanisms of how artificial intelligence fueled and shaped political and social action.

> Our vocation as historians… is vital for understanding just the contingency of the entire array of structures—physical, intellectual, social—upon which AI has been built.

As historians, we could treat large-scale automatic decision systems entirely at a remove, as unrelated to, or even a threat to, our largely humanistic vocation. Historians will be limiting themselves if they refuse to draw critically and judiciously upon the range of digital methods focused on empirical specificity—the techniques central to new machine learning/AI—to contend with the transformation of scale and new forms of communication, solidarity, and organization mediated by large-scale algorithmic systems. [24] Indeed, as historians we should play a fundamental role in assuring a critical use of such tools in thinking through the past—and the present. My claim here is that the history of data and AI systems is crucial to doing such history using data and machine learning reflectively. Solutions to past epistemological crises in history from the late seventeenth century onward meant rejecting naive positivism in the name of stressing the critical use of evidence, from the documentary to the numismatic to the biological. In his history of climate modeling, Paul Edwards describes the practice of "infrastructural inversion" in science: "data aren't data until you have turned the infrastructure upside down to find out how it works."[25] We historians too need to turn infrastructures of data collection upside down, using technical histories of data and algorithms, to be sure, but ranging far beyond, to understand the creation and analysis of data in the contexts of its production and its use and to track the contestation and consolidation of those infrastructures in the organization of social, intellectual, and cultural life.

The best humanistic scholarship using digital tools, including data-driven machine learning tools, illustrates the intertwining of the hermeneutic with critical digital praxis.[26] Reflecting on the scholarship using the Trans-Atlantic Slave Trade Database, Jessica Marie Johnson emphasizes the labor required to use the data produced through the "transmutation of black flesh into integers and fractions." To gain any historical access to the "lives of women and children from the data set" she argues, "required a methodology attuned to black life and to dismantling the methods used to create the manifests in the first place, then designing and launching an interface responsive to the desire of descendants of slaves for reparation and redress."[27] In her critical practice, Johnson insists on interpretation based on swathes of contextual knowledge as a precondition for then using computational data.

Recent work on the biases, politics, and ethics of data sets calls for data scientists to draw upon the commitments of competencies of archivists to address fundamental questions around data curation, from consent to documentation.[28] Large-scale language models, for example, testify to just how historical automatic decision systems are: biases omnipresent in large text databases ramify through systems designed precisely to capture the full, often horrific, richness of that data, but prone to accept and overemphasize dated if dominant approaches to the world and to other people, what some key researchers call "the tendency of training data ingested from the Internet to encode hegemonic worldviews."[29] Reading sources for hegemonic and antihegemonic worldviews in rich contexts is just what historians at their best do—and what our communities have to offer a broad community of other committed scholars and activists.

We might best conceive the long-term role for digital methods, including machine learning or AI systems, as additions to the auxiliary sciences of history—those technical domains of diplomatics, numismatics, chronology—that enable without determining historical labor—indeed, enable reflection upon the very conditions of the construction of our archives and our historical practices.[30] Some years ago Michel de Certeau noted the danger of insulating historical practice from technique: "Insofar as the university is foreign to practice and to technicalities, everything that places history in rapport with techniques is classified as 'auxiliary science.'" In the past this meant "epigraphy, papyrology, paleography, diplomatics, codicology" and "today, musicology, 'folklorism,' computer science." History proper, Certeau noted "would only begin with the 'noble speech' of interpretation. It would finally be an art of discourse delicately erasing all traces of labor."[31] In our practice, we should do the opposite: reveal the traces of labor, whether in doing a history of artificial intelligence or telling it, that enable our hermeneutic reflective work. Easy and old binaries will only disempower us and allow simplistic stories to prevail—be they techno-determinist, techno-utopian, techno-dystopian, or neoliberal—rather than allowing the nuanced, fraught, contradictory histories of diverse societies mediated and directed though algorithmic decision-making systems to enable deeper critical engagement. The new artificial intelligence emerged from the desire of capturing a granular complex reality; our best histories must be able to capture that granular complex reality of artificial intelligence.

Matthew L. Jones is the Smith Family Professor of History at Princeton University where he focuses on the history of science and technology in early modern Europe and on recent information technologies. With his collaborator Chris Wiggins, he published in 2023 *How Data Happened*, a history of the science, politics, and power of data, statistics, and machine learning from the 1800s to the present.

1 Donald Michie, "Machine Intelligence Research Unit l: Main Research Theme: Computer Induction" [11.1982], British Library Add. Ms. 88975/5/2. Emphasis in the original.

2 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.

3 I use the terms "AI," "automated decision systems," and "machine learning" more or less interchangeably throughout.

4 Stephanie Dick, "Artificial Intelligence," *Harvard Data Science Review* 1, no. 1 (July 1, 2019), https://doi.org/10.1162/99608f92.92fe150c; Margaret A. Boden, "GOFAI," in *The Cambridge Handbook of Artificial Intelligence*, ed. Keith Frankish and William M. Ramsey (Cambridge, 2014), 89–107, https://doi.org/10.1017/CBO9781139046855.007.

5 Dominique Cardon, Jean-Philippe Cointet, and Antoine Mazières, "La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle," *Réseaux* 211, no. 5 (2018): 173–220, https://doi.org/10.3917/res.211.0173.

6 I've offered a schematic approach to this history in Matthew L. Jones, "How We Became Instrumentalists (Again): Data Positivism since World War II," *Historical Studies in the Natural Sciences* 48, no. 5 (November 2018): 673–84, https://doi.org/10.1525/hsns.2018.48.5.673.

7 Hanna Wallach, "Big Data, Machine Learning, and the Social Sciences," *Medium*, December 23, 2014, https://medium.com/@hannawallach/big-data-machine-learning-and-the-social-sciences-927a8e20460d.

8 Margaret A. Boden, *Mind as Machine: A History of Cognitive Science* (New York, 2006); Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge, 2010); Roberto Cordeschi, *The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics* (Dordrecht, 2002); In science studies more broadly, crucial studies include D. E. Forsythe, "Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence," *Social Studies of Science* 23, no. 3 (August 1993): 445–77, https://doi.org/10.1177/0306312793023003002; Lucy A. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions*, 2nd ed. (Cambridge, 2007). More recent studies include Stephanie Dick, "Of Models and Machines: Implementing Bounded Rationality," *Isis* 106, no. 3 (2015): 623–34; Joseph Adam November, *Biomedical Computing: Digitizing Life in the United States* (Baltimore, 2012); Sabina Leonelli, *Data-Centric Biology: A Philosophical Study* (Chicago, 2016); Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World* (Cambridge, MA, 2018); Cardon, Cointet, and Mazières, "La revanche des neurones"; Yarden Katz, *Artificial Whiteness: Politics and Ideology in Artificial Intelligence* (New York, 2020), chaps. 1–2; Aaron Mendon-Plasek, "Mechanized Significance and Machine Learning: Why It Became Thinkable and Preferable to Teach Machines to Judge the World," in *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*, ed. Jonathan Roberge and Michael Castelle (Cham, 2021), 31–78, https://doi.org/10.1007/978-3-030-56286-1_2; Jonathan Penn, "Inventing Intelligence: On the History of Complex Information Processing and Artificial Intelligence in the United States in the Mid-Twentieth Century" (Ph.D. thesis, University of Cambridge, 2021), https://doi.org/10.17863/CAM.63087.

9 Key examples include Eden Medina, *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile* (Cambridge, MA, 2011); Colin Garvey, "Artificial Intelligence and Japan's Fifth Generation: The Information Society, Neoliberalism, and Alternative Modernities," *Pacific Historical Review* 88, no. 4 (November 2019): 619–58, https://doi.org/10.1525/phr.2019.88.4.619; Ekaterina Babintseva, "From Pedagogical Computing to Knowledge-Engineering: The Origins and Applications of Lev Landa's Algo-Heuristic Theory," in *Abstractions and Embodiments: New Histories of Computing and Society*, ed. Stephanie Dick and Janet Abbate (Baltimore, 2022.); Sandeep Mertia, ed., *Lives of Data: Essays*

on Computational Cultures from India (Amsterdam, 2020).

10 I was fortunate to be a coorganizer of a Mellon-Sawyer seminar on the History of Artificial Intelligence, based in Cambridge, England, aimed precisely at this effort. The volume of papers to appear soon will dramatically augment our knowledge of the global history of AI. https://www.hps.cam.ac.uk/research/projects/histories-of-ai. For the commercialization of some earlier AI systems, see Hallam Stevens, "The Business Machine in Biology—The Commercialization of AI in the Life Sciences," *IEEE Annals of the History of Computing* 44, no. 01 (January 2022): 8–19, https://doi.org/10.1109/MAHC.2021.3104868.

11 Sarah Elizabeth Igo, *The Known Citizen: A History of Privacy in Modern America* (Cambridge, MA Harvard, 2018).

12 Ruha Benjamin, *Race after Technology: Abolitionist Tools for the New Jim Code* (Cambridge, 2019); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York, 2017).

13 Khalil Gibran Muhammad, *The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America* (Cambridge, MA, 2010); Daniel B. Bouk, *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual* (London, 2015); Katz, *Artificial Whiteness*.

14 Paul N. Edwards, Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman, "Science Friction: Data, Metadata, and Collaboration," *Social Studies of Science* 41, no. 5 (2011): 667–90; Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out* (Cambridge, MA, 1999).

15 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, 2016), 311–28.

16 Nicole Starosielski, *Fixed Flow: Undersea Cables as Media Infrastructure* (Urbana, IL, 2015).

17 Mél Hogan, "Data Flows and Water Woes: The Utah Data Center," *Big Data & Society* 2, no. 2 (December 2015), https://doi.org/10.1177/2053951715592429; Nathan Ensmenger, "The Environmental History of Computing," *Technology and Culture* 59, no. 4 (2018): S7–33, https://doi.org/10.1353/tech.2018.0148; Zane Griffin Talley Cooper, "Of Dog Kennels, Magnets, and Hard Drives: Dealing with Big Data Peripheries," *Big Data & Society* 8, no. 2 (July 2021), https://doi.org/10.1177/20539517211015430.

18 Margaret O'Mara, *The Code: Silicon Valley and the Remaking of America* (New York, 2019).

19 Sarah T. Roberts, *Behind the Screen: Content Moderation in the Shadows of Social Media* (New Haven, 2019); Lilly Irani, "Justice for 'Data Janitors,'" *Public Books* (blog), January 15, 2015, http://www.publicbooks.org/justice-for-data-janitors/; Mary L. Gray and Siddharth Suri, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass* (Boston, 2019).

20 Joanna Radin, "'Digital Natives': How Medical and Indigenous Histories Matter for Big Data," *Osiris* 32, no. 1 (September 2017): 43–64, https://doi.org/10.1086/693853; Marlee Tichenor, "The Power of Data: Global Malaria Governance and the Senegalese Data Retention Strike," in *Metrics: What Counts in Global Health*, ed. Vincanne Adams (Durham, NC, 2016), 105–24, https://doi.org/10.1215/9780822374480-005.

21 Leslie Willcocks, "Robo-Apocalypse Cancelled? Reframing the Automation and Future of Work Debate," *Journal of Information Technology* 35, no. 4 (December 2020): 286–302, https://doi.org/10.1177/0268396220925830.

22 Alexander Campolo, "Signs and Sight: Jacques Bertin and the Visual Language of Structuralism," *Grey Room* 78 (February 2020): 34–65, https://doi.org/10.1162/grey_a_00286.

23 For the urgency of parallel questions, see Jo Guldi, "Working on a Deadline," *Isis* (forthcoming).

24 For an important statement on the limits created by the infrastructural demands that would be required for digital history at scale, see Ben Schmidt, "Two Volumes: The Lessons of Time on the Cross," Ben Schmidt (blog), December 5, 2019, https://benschmidt.org/post/2019-12-05-totc/2019-aha/.

25 Paul Edwards, *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming* (Cambridge, MA, 2010), 20. See also Alan Liu, "Toward Critical Infrastructure Studies: Digital Humanities, New Media Studies, and the Culture of Infrastructure," (lecture, University of Connecticut, Storrs, CN, Feb. 23, 2018), https://liu.english.ucsb.edu/toward-critical-infrastructure-studies-u-connecticut/.

26 Catherine D'Ignazio and Lauren F. Klein, *Data Feminism* (Cambridge, MA, 2020); Richard Jean So, Hoyt Long, and Yuancheng Zhu, "Race, Writing, and Computation: Racial Difference and the US Novel, 1880–2000," *Journal of Cultural Analytics* 3, no. 2 (2018), https://doi.org/10.22148/16.031. Jo Guldi, "Critical Search: A Procedure for Guided Reading in Large-Scale Textual Corpora," *Journal of Cultural Analytics* 3, no. 1 (2018), https://doi.org/10.22148/16.030.

27 Johnson, “Markup Bodies,” 65. See the discussion of this paper in Schmidt, “Two Volumes.”

28 Eun Seo Jo and Timnit Gebru, “Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning,” in *FAT* ‘20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, 2020), 306–16, https://doi.org/10.1145/3351095.3372829.

29 Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Margaret Mitchell, “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?,” in *FAccT ‘21: Proceedings of the 2021 ACM Conference on Fairness, Accountability and Transparency* (New York, 2021), 610–23, at 616.

30 Compare, Lara Putnam, “The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast,” *American Historical Review* 121, no. 2 (April 2016): 377–402, https://doi.org/10.1093/ahr/121.2.377.

31 Michel de Certeau, *The Writing of History*, trans. Tom Conley (New York, 1988), 69.