

Making Automation Work for Social Scientists

M. J. Crockett

Many scientists of all stripes are optimistic about new research opportunities afforded by “AI.” This term refers to a diverse set of automated and semiautomated technologies including (but not limited to) predictive machine learning systems, automated transcription, generative models that produce images and text, and chatbots built on large language models (LLMs) that simulate conversation. Within social science, researchers envision using these technologies to accelerate and expand their abilities to gather, analyze, and simulate human data.¹ Some have even proposed automating the entire research pipeline, with LLMs replacing human subjects and scientists.²

I’ve watched these developments with a mixture of excitement and alarm. My lab at Princeton University has used predictive machine learning to analyze public discourse on social media, and recent advances in LLMs enable us to pursue research questions that were out of reach just a few years ago.³ At the same time, I worry that widespread adoption of automation across the research pipeline is outpacing thoughtful consideration of its many risks and harms, including harms to us as knowledge workers.⁴ To be clear: my worries don’t reflect a nostalgia for some idyllic analog past. Rather, I’m concerned with what scientific futures we foreclose if we cede our epistemic agency to automated systems.

I’ve begun to find a way through this minefield in conversation with my brilliant and generous collaborator Lisa Messeri, an anthropologist of science and technology who investigates how expert communities create new fields of inquiry and innovation.⁵ Lisa introduced me to scholarship illuminating how political, economic, and cultural forces shape our collective imagination about what technology is and can be. Here, I offer some insights that have emerged from our collaboration and shape the way I approach research in my lab, with the aim of preserving our epistemic agency under intense pressure to give it away.⁶

First, industry marketing can mislead us into thinking that new technologies are more capable than they actually are. Technology companies are selling AI with overhyped claims of superhuman or even magical abilities.⁷ Some sci-

entists imagine LLMs as oracles, capable of extracting “objective” truths from the published literature or large datasets.⁸ Lisa and I name this imaginary as an *illusion of objectivity*: LLMs reflect the viewpoints of their engineers and those represented in their training data.⁹ Rather than removing bias from the scientific process, this technology entrenches dominant viewpoints while simultaneously obscuring them from view.¹⁰ Just as social scientists have begun to make progress in diversifying the questions we ask and the people we study, wholesale adoption of automation threatens to reverse that progress.

When automated tools are controlled by private industry, scientists cannot access information crucial for vetting the capabilities of these tools. Accounting for biases in models is challenging because of their opacity. This is especially the case for proprietary LLMs like the GPT product line, because OpenAI has not disclosed the contents of its models’ training data. Moreover, the outputs of proprietary LLMs vary unpredictably over time, as engineers tweak model parameters behind the scenes to satisfy corporate goals.¹¹ We cannot build a robust and reproducible social science using models whose construction is opaque, whose outputs are unstable, and whose control rests in the hands of CEOs with very different goals than ours.¹²

Social scientists are increasingly seeking financial support from the technology industry, especially as public funding becomes scarce. This has at least two consequences for our science. First, we are incentivized to make our research questions fit the capabilities of automated tools. But not all questions are amenable to computational analysis.¹³ We suffer from an *illusion of exploratory breadth* when we mistake the subset of questions automated tools can address with the broader set of questions we can ask about the social world.¹⁴ Second, industry funding encourages scientists to focus on questions that are friendly to technology companies.¹⁵ These complementary forces can lead to the development of scientific monocultures, where a narrow set of methods and questions dominates knowledge production, making science less innovative and robust.¹⁶

The varied risks of automation for social science all flow from the outsourcing of scientific judgment to automated systems and the companies that control them. Recognizing this suggests concrete steps we can take to reclaim our epistemic agency in the short term and preserve it in the long term. Instead of being seduced by fantasies of “superintelligent” general purpose systems that can replace the work of scientists, we can build automated tools for specific scientific tasks. Rather than relying on opaque proprietary models, we can insist on using open-source models with transparent documentation.¹⁷ To prevent the formation of scientific monocultures, we can recognize that qualitative and ethnographic work offers distinctive ways of understanding those aspects of social life that resist quantification, and we can advocate for continued investment in these diverse methodologies.¹⁸ We should also require our colleagues to be more transparent about conflicts of interest that arise from industry funding, encouraging industry-independent research instead.¹⁹

These things called “AI” will surely change in the years to come, but this does not mean current insights have an expiration date. Our colleagues in the humanities and humanistic social sciences have long recognized that technological advances rarely create entirely new dilemmas, instead reproducing old dilemmas in new forms. The communal values of science have always been in tension with private commercial interests.²⁰ Now, by aiming to commodify knowledge production itself, the technology industry makes this tension harder to ignore. Our urgent task at present is to question the uncritical embrace of artificial intelligence and articulate a future social science in which automation works for us.

AUTHOR’S NOTE

Thanks to Emily Bender, Lisa Messeri, and Alondra Nelson for helpful comments.

ABOUT THE AUTHOR

M. J. Crockett is Director of the Future Values Initiative at the Princeton University Center for Human Values and Professor in the Department of Psychology at Princeton, where they also direct the Crockett Lab. They have published in such journals as *Science*, *Nature*, and *Trends in Cognitive Sciences*.

ENDNOTES

- ¹ Igor Grossmann, Matthew Feinberg, Dawn C. Parker, et al., “AI and the Transformation of Social Science Research,” *Science* 380 (6650) (2023): 1108–1109; and Christopher A. Bail, “Can Generative AI Improve Social Science?” *Proceedings of the National Academy of Sciences* 121 (21) (2024): e2314021121.
- ² Benjamin S. Manning, Kehang Zhu, and John J. Horton, “Automated Social Science: Language Models as Scientist and Subjects,” NBER Working Paper 32381 (National Bureau of Economic Research, 2024); and Sebastian Musslick, Younes Strittmatter, and Marina Dubova, “Closed-Loop Scientific Discovery in the Behavioral Sciences,” PsyArXiv (2024), <https://doi.org/10.31234/osf.io/c2ytb>.
- ³ William J. Brady, Killian McLoughlin, Tuan N. Doan, and Molly J. Crockett, “How Social Learning Amplifies Moral Outrage Expression in Online Social Networks,” *Science Advances* 7 (33) (2021): eabe5641; Killian L. McLoughlin, William J. Brady, Aden Goolsbee, et al., “Misinformation Exploits Outrage to Spread Online,” *Science* 386 (6725) (2024): 991–996; and William J. Brady, Killian L. McLoughlin, Mark P. Torres, et al., “Overperception of Moral Outrage in Online Social Networks Inflates Beliefs About Intergroup Hostility,” *Nature Human Behavior* 7 (6) (2023): 917–927.

- ⁴ Laura Weidinger, John Mellor, Maribeth Rauh, et al., “Ethical and Social Risks of Harm from Language Models,” arXiv (2021), <https://doi.org/10.48550/arXiv.2112.04359>; Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, “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* (Association for Computing Machinery, 2021), 610–623; Karen Hao, *Empire of AI: Dreams and Nightmares in Sam Altman’s OpenAI* (Penguin Press, 2025); and Committee on Artificial Intelligence and Academic Professions, *Artificial Intelligence and Academic Professions* (American Association of University Professors, 2025), <https://www.aaup.org/reports-publications/aaup-policies-reports/topical-reports/artificial-intelligence-and-academic>.
- ⁵ Lisa Messeri, *Placing Outer Space: An Earthly Ethnography of Other Worlds* (Duke University Press, 2016); and Lisa Messeri, *In the Land of the Unreal: Virtual and Other Realities in Los Angeles* (Duke University Press, 2024).
- ⁶ Molly J. Crockett, Xuechunzi Bai, Sayash Kapoor, et al., “The Limitations of Machine Learning Models for Predicting Scientific Replicability,” *Proceedings of the National Academy of Sciences* 120 (33) (2023): e2307596120; Lisa Messeri and Molly J. Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research,” *Nature* 627 (8002) (2024): 49–58; and Molly J. Crockett and Lisa Messeri, “AI Surrogates and Illusions of Generalizability in Cognitive Science,” *Trends in Cognitive Sciences* 30 (3) (2025): 203–215, <https://doi.org/10.1016/j.tics.2025.09.012>.
- ⁷ Emily M. Bender and Alex Hanna, *The AI Con: How to Fight Big Tech’s Hype and Create the Future We Want* (HarperCollins, 2025).
- ⁸ Messeri and Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research”; and Molly J. Crockett, “Modern Maxims for an AI Oracle,” *Nature Machine Intelligence* 7 (1) (2025): 4–5.
- ⁹ Bender, Gebru, McMillan-Major, and Shmitchell, “On the Dangers of Stochastic Parrots”; and Messeri and Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research.”
- ¹⁰ Crockett and Messeri, “AI Surrogates and Illusions of Generalizability in Cognitive Science.”
- ¹¹ Christopher Barrie, Alexis Palmer, and Arthur Spirling, “Replication for Language Models: Problems, Principles, and Best Practices for Political Science,” working paper, https://arthurspirling.org/documents/BarriePalmerSpirling_TrustMeBro.pdf (accessed March 17, 2026).
- ¹² Robert K. Merton, “The Normative Structure of Science,” in *The Sociology of Science: Theoretical and Empirical Investigations*, ed. Norman W. Storer (University of Chicago Press, 1979), 267–278.
- ¹³ danah boyd and Kate Crawford, “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon,” *Information, Communication & Society* 15 (5) (2012): 662–679.
- ¹⁴ Messeri and Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research.”
- ¹⁵ Joseph Bak-Coleman, Cailin O’Connor, Carl Bergstrom, and Jevin West, “The Risks of Industry Influence in Tech Research,” arXiv (2025), <https://doi.org/10.48550/arXiv.2510.19894>; and Meredith Whittaker, “The Steep Cost of Capture,” *Interactions* 28 (6) (2021): 50–55.
- ¹⁶ Messeri and Crockett, “Artificial Intelligence and Illusions of Understanding in Scientific Research”; Joseph D. O’Brien, Hannah Rubin, Kekoa Wong, and Sina Fazelpour,

- “Epistemic Monocultures and the Effect of AI Personalization,” *Proceedings of the Annual Meeting of the Cognitive Science Society* 47 (2025); and Qianyue Hao, Fengli Xu, Yong Li, and James Evans, “Artificial Intelligence Tools Expand Scientists’ Impact but Contract Science’s Focus,” arXiv (2024), <https://doi.org/10.48550/arXiv.2412.07727>.
- ¹⁷ Bail, “Can Generative AI Improve Social Science?”; Michael C. Frank, “Openly Accessible LLMs Can Help Us to Understand Human Cognition,” *Nature Human Behavior* 7 (11) (2023): 1825–1827; and “European Open-Source AI Index,” Centre of Language and Speech Technology, Radboud University, <https://osai-index.eu/the-index> (accessed March 17, 2026).
- ¹⁸ boyd and Crawford, “Critical Questions for Big Data”; and Lauren Klein, Meredith Martin, André Brock, et al., “Provocations from the Humanities for Generative AI Research,” arXiv (2025), <https://doi.org/10.48550/arXiv.2502.19190>.
- ¹⁹ Bak-Coleman, O’Connor, Bergstrom, and West, “The Risks of Industry Influence in Tech Research”; and Janet Vertesi and J. Nathan Matias, “Divesting from Big Tech: Alternative Possibilities for Research and Futuring in Social Computing,” in *CSCW ’23 Companion: Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*, ed. Casey Fiesler, Loren Terveen, Morgan Ames, et al. (Association for Computing Machinery, 2023), 401–404, <https://doi.org/10.1145/3584931.3608436>.
- ²⁰ Merton, “The Normative Structure of Science.”