PNAS PERSPECTIVE OPEN ACCESS

Can Generative AI improve social science?

Christopher A. Bail[a,b,c,1]

Edited by David Lazer, Northeastern University, Boston, MA; received September 7, 2023; accepted April 5, 2024, by Editorial Board Member Mark Granovetter

for research with Generative AI, prevent these tools from reproducing the academic caste system, and allow social scientists to develop solutions to future challenges and prevent these tools from being repurposed for malicious purposes.

Several caveats are in order. First, my analysis is limited to social science and thus does not engage with the many different ways Generative AI might shape other fields. Second, I focus on the impact of Generative AI on scientific research, and not its broader impact on social life—a topic

people in creative industries (1). Models such as DALL-E and Stable Diffusion create such visual content through text prompts—searching for connections between patterns in the co-occurrence of words and the arrangement of pixels—that allow a user to request highly specialized visual content.

Opportunities for Social Science with Generative AI

example, makes them respond to public opinion surveys in a manner that is very similar to real respondents with the same attributes. Some argue such "silicon samples" could be used to produce more diverse samples than the convenience samples utilized by so many university researchers—and may also allow researchers to administer lengthier survey instruments, since LLMs have potentially unlimited attention spans (12). At the same time, more recent research indicates GPT 3.5 turbo produces accurate mean estimates of

in the 2000s and 2010s. But until recently, these chatbots also appeared incapable of passing the Turing test, since they struggled to generate original content and frequently redirected conversations—or failed to follow other conventions in human conversation that made them fairly easy to identify. Generative AI holds the potential to create more realistic human-like interactions given that many such tools are trained on larger amounts of data that describe human interactions—and also because of recent technical innovawhere respondents were recruited to interact with LLMs that were prompted to impersonate members of the opposing political party for ten minutes. Though respondents were told they might interact with automated accounts during the study's informed consent dialog, most participants expressed uncertainty about whether they interacted with humans or bots. These findings are preliminary due to the study's small sample size, but the research design indicates LLMs may be useful for conducting research on group-level processes pro-

of agents, it provides a proof of concept that Generative AI has the potential to advance social simulation research.

More recent studies indicate LLMs can be integrated within ABMs to develop or test more sophisticated theories of human behavior. For example, Törnberg et al. create a simulated social media platform with five hundred agents whose behaviors are calibrated using data from the American National Election Study (ANES) (39). The agents are prompted to read news stories and make posts—or like difficult to study may not be well represented in the training data used to create Generative AI.

Can Generative AI Improve Text Analysis? Regardless of whether Generative AI can effectively simulate human behavior, it may also help social scientists with other common research tasks such as content analysis of text-based data. Wu et al. demonstrate GPT-3.5 can produce accurate

coders are also prone to a variety of well-documented errors that range from subjective bias to inconsistency and lack of attention—particularly when researchers organize small teams to code documents in a coordinated fashion. LLMs can also exert bias and be inconsistent, as I discuss in further detail below. But LLMs may enable social scientists to examine corpora of unprecedented size with unforeseen speed. Rather than taking a random sample of documents, for example, social scientists now have the potential to code

Generative AI models—such as GPT-4—are largely unknown. Without access to the types of training data fed into such models, researchers can only examine "known unknowns." If poor elderly people in rural areas are unable to voice their collective concern about how Generative AI represents them, for example, researchers may be unlikely to identify such bias.

A key question for social scientists is whether the bias of Generative AI is a "bug" or a "feature" for research purposes. We often design experiments that examine the impact of

developers use to help each other write code. As enthusiasm about the capacity of Generative AI to write code peaked, some users created bots that automatically passed people's questions about software to LLMs. Though many of the answers produced by the LLM were high quality, others were completely incorrect. The website quickly announced a new policy banning LLMs to prevent a situation where users would struggle to distinguish the good information from the bad.

Researchers who rely upon LLMs to perform literature recreates the risk that an AI agent could encourage conflict between human participants. Some of these risks might be mitigated via content moderation filters that are currently available for some LLMs—and through rigorous testing of the prompts used to guide LLMs in research settings. Yet given the probabilistic nature of these models—and the ever-changing ways abuse and harassment can occur in online settings—such strategies will require great care.

Another strategy is to design studies where Generative AI

weighed against the efficiencies they create, however. One study, for example, suggests the carbon emissions of writing and illustrating are lower for AI than for humans (71).

Is Research with Generative AI Replicable? A key pillar of the open-science movement is that researchers should design studies that can be replicated by others. Though Generative AI may help researchers increase the external validity of

Creating Open-Source Infrastructure for Social Science Research

As the sections above describe, Generative AI has many limitations for social science research—most of which we are only beginning to understand. How can social scientists work together to minimize the risks of research with Generative AI without sacrificing the many opportunities it creates? Accomplishing this agenda will require deeper understanding

have privacy benefits. Prompts used by researchers could be carefully protected, instead of being potentially resold to third-parties or used to develop future models (72).

Open-source models also often create and sustain a community of people with shared concerns. Rather than guessing when and how proprietary models may exhibit bias—or endlessly testing different prompts to achieve research goals—social scientists could work together to identify the limitations of Generative AI tools for social science

Such an organization could also explore broader common goods, such as the creation of a large silicon sample of human populations that researchers can use to conduct preliminary tests of human subjects, or an open-source codebase for integrating LLMs into agent-based models.

Conclusion

Few technologies have created so much excitement—and

the future of AI research will require training models to better understand the science of social relationships—for example, how an AI agent should interact in group settings where the goal is not simply to provide utility for a single user, but to navigate the more complex challenges associated with emergent group behaviors. If I am correct, social scientists may soon find themselves at the center of efforts to "reverse engineer" what the sociologist William H. Sewell Jr. calls the "social sense." That is, the ability for Generative AI to

43. Y. Li, Y. Zhang, L. Sun, Metaagents: Simulating interactions of human behaviors for LLM-based task-oriented coordination via collaborative generative agents. arXiv [Preprint] (2023). https://arxiv.org/abs/2310.06500 (Accessed 19 December 2023).
44. N. Ghaffarzadegan, A. Majumdar, R. Williams, N. Hosseinichimeh, Generative agent-based modeling: Unveiling social system dynamics through coupling mechanistic models with generative artificial intelligence. arXiv [Preprint] (2023). https://arxiv.org/abs/2309.11456 (Accessed 19 December 2023).
45. B. Xiao, Z. Yin, Z. Shan, Simulating public administration crisis: A novel generative agent-based simulation system to lower technology barriers in social science research. arXiv [Preprint] (2023). https://arxiv.org/abs/2311.06957 (Accessed 19 December 2023).
46. A. S. Vezhnevets *et al*., Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia. arXiv [Preprint] (2023). https://arxiv.org/abs/2312.03664 (Accessed 19 December 2023).
47. Z. Kaiya *et al*., Lyfe agents: Generative agents for low-cost real-time social interactions. arXiv [Preprint] (2023). https://arxiv.org/abs/2310.02172 (Accessed 19 December 2023).
48. R. Axtell, R. Axelrod, J. M. Epstein, M. D. Cohen, Aligning simulation models: A case study and results. *Comput. Math. Organ. Theory* **1**, 123–141 (1996).
49. P. Y. Wu, J. Nagler, J. A. Tucker, S. Messing, Large language models can be used to scale the ideologies of politicians in a zero-shot learning setting. arXiv [Preprint] (2023). https://doi.org/10.48550/arXiv.2303.12057 (Accessed 20 December 2023).