

Science communication with generative AI



Generative AI tools can quickly translate or summarize large volumes of complex information. This technology could revolutionize the way that we communicate science, but there are many reasons for caution. We asked six experts about the potential and pitfalls of generative AI for science communication.

How can generative AI tools help us to communicate science?

Shirley S. Ho: Generative AI tools have introduced new synergy to science communication by potentially generating content at an easier and faster pace than before. For example, tools such as [ChatGPT4](#) and [txyz.ai](#) can enable science communicators to upload scientific papers and translate scientific jargon to make it more understandable for laypeople. These tools can produce a large volume of content quickly. The introduction of generative AI tools enhances the ease with which scientific jargon can be translated, as well as the speed of content creation for science communication.

Aylin Caliskan & Jevin West: Generative AI presents new opportunities for human–AI collaboration in science communication. There is reason to be excited. For example, generative AI tools such as ChatGPT and Semantic Scholar's [TLDR](#) feature are competent at summarization and paraphrasing. Generative AI can spark the frozen writer or creator. One can ask for an introduction to a particular topic using ChatGPT or other language models, or they can take a collection of notes or bullet points and turn them into readable sentences. One play around with the image features to explore different ways of building a visual for explaining a new model or for presenting a slide for a talk at a conference. Generative AI-powered translation tools such as ChatGPT can help non-native English speakers to translate science – although the jury is out as to whether existing, more-customized models for translation and grammar-checking (such as Google Translate, Bing Translator, DeepL or Grammarly) are better, when the cost to the environment is considered. And generative AI might aid the overtaxed peer-review system, but only if humans remain in the loop.



Generative AI could potentially help reviewers to identify inconsistencies in results and conclusions, and could summarize human reviewer notes.

Lisa Messeri & M. J. Crockett: Scientists are anticipating how generative AI tools can augment or replace human labour across the research pipeline¹, from acting as an Oracle that can search, evaluate and summarize scientific literatures² to an Arbiter that can evaluate scientific merit and replicability of findings³. The presumed savings to time and cost that make generative AI appealing when producing scientific knowledge⁴ similarly makes it a tempting solution to the challenge of communicating scientific findings – both to other scientists as well as to the broader public. This hopeful vision of AI as Scribe posits that generative AI could help to translate research findings for different audiences and ease the burden on already overtasked researchers.

What are your concerns about generative AI tools and science communication?

Amanda Alvarez: Given the much-publicized propensity of generative AI tools to produce

nonsense, science communicators should consider whether generative AI is in fact completely antithetical to the very purpose of their work.

A.C. & J.W.: If the goal of science were to write papers, generative AI would be a watershed moment. But this is not the goal. Science is a social process with the goal of understanding the world in which we live. Writing more papers is not the limiting resource; we must proceed with caution so that we advance human knowledge, rather than just produce piles of papers.

Precision is a key feature of science communication, but is not a strong suite of generative AI. Hallucinations and 'hallu-citations' are an intrinsic property of generative AI. We need to be aware of these baseline errors across domains and not underestimate the effort that is required to fact-check the output of generative AI. In addition, generative AI converges on stylistic patterns from its training data that lose important nuance and therefore precision.

As with most machine technology that is reliant on large, human data, biases exist. ChatGPT has been shown to perpetuate gender biases in machine translation from

one language to another⁵. When using AI for science communication, we must be aware of these biases and do our best not to perpetuate them.

Biases also affect the science that gets highlighted in generative AI-mediated search and field-level summaries. Hype in the training data – from media articles and social media – can lead to an over-reliance on select studies and authors.

L.M. & M.J.C.: Using generative AI as a tool for communicating science could eliminate diversity from the pool of science communicators. Decades of theoretical and empirical work recognizes that diversity (including lived experiences and academic training) strengthens scientific knowledge production^{6,7}. Likewise, diverse communicators of science will highlight different meanings, limitations and implications of scientific work. Tasking generative AI with science communication threatens to erase diverse interpretations of scientific work⁸, which narrows the potential for science to affect the broader public.

In place of diverse science communicators, what perspective does generative AI instead offer? Many envision generative AI as providing a neutral or universal standpoint that desirably eliminates bias from scientific communication. This vision clashes with the fact that datasets used to train generative AI tools such as ChatGPT reflect the hegemonic perspectives of English-speaking people with internet access, who tend to skew younger and who are willing to publicly express their views^{9,10}. Thus, using generative AI tools for science communication not only erases diversity but also overrepresents the perspectives of those who have shaped the production and application of science for centuries. Embedding those perspectives in computer code makes them less visible, but not less pervasive.

A.A.: Do we as science communicators and purveyors of science – a uniquely human endeavour – really want to forfeit that wonder and responsibility to the likes of a faceless chatbot? I expect that level of inauthenticity from a bank or an online retailer; do I also want to set that expectation for universities or trusted news outlets?

How is trust in generative AI tools a concern when using them for science communication?

A.A.: The explosion of generative AI has stirred up questions of legitimacy and

trustworthiness, and this is evidenced by the Merriam-Webster dictionary's choice for the word of the year in 2023: **authentic**. The science communication field faces challenges from lightning-quick generative AI systems that can fabricate references, boost misinformation and pollute the knowledge pool itself.

S.S.H.: AI-generated content may not always be accurate, reliable or transparent, which is crucial in science communication. Deploying AI to generate content at scale raises concerns about the potential for the large-scale creation and spread of scientific misinformation. Besides, the originality and integrity of scientific research articles are in doubt if researchers do not disclose the use of generative AI¹¹.

Against this backdrop, do the lay public trust science content curated by science communicators, with the assistance of generative AI tools? Do the lay public trust the technology companies and developers of generative AI, given the fact that many are developing the technology for profit instead of for the good of humanity?

As key consumers of science information, the lay public often rely on heuristic cues such as trust to assess the credibility of scientific information¹² and to make decisions about science^{13,14}. Studies have underscored the importance of trust in shaping public attitudes towards and intention to use generative AI tools¹⁵. Recent research has also shown that trust in generative AI tools can be built through enhancing its transparency, fairness, accountability and explainability¹⁵, although this can be an issue because the process remains opaque to many. As a fundamental imperative of science communication is to establish trust in science, a lack of trust in generative AI tools in producing science content may become a real detractor to this goal.

This is exacerbated by a situation in which many science communicators are not sufficiently trained to use generative AI tools to develop content, despite having experimented with the new technology. To gain the trust of the public, science communicators need to equip themselves with the skillsets to verify and crosscheck the content that they produce with the aid of generative AI tools. Companies should also develop policies that make declaration of the use of generative AI tools for science communication mandatory.

Evidently, developing generative AI goes beyond trustworthy AI; trustworthy science communicators are also part of the technological ecosystem. Can the use of generative AI tools by science communicators measure

up to public expectations of trustworthiness? This is an important question to ponder as the interaction between science communicators and the next generation of generative AI co-evolves.

Are generative AI tools compatible with culturally specific or sensitive science communication?

A.A.: ChatGPT, DALL-E and other generative AI tools have set us up for an arms race to produce ever more volumes of eye-catching, interactive and multimodal media. But for whom? Audience (read: humans) is key here.

Good science communication is a dialogue, but generative AI tools such as ChatGPT do not know who they are talking to or why. In certain situations, bot-to-human communication is the wrong approach. For example, reporting on #MeToo in academia requires a sharp grasp of facts, legal considerations, timelines, confidentiality and sensitivity. The debate around contentious and culturally sensitive scientific infrastructure such as the **Mauna Kea Observatories** warrants human empathy and historical awareness. ChatGPT obviously cannot conduct interviews or investigative journalism, and the sourcing of its facts is unknown given the opaque training data. Meaning is more than just words credibly strung together.

“Good science communication is a dialogue, but generative AI tools ... do not know who they are talking to or why”

Similarly, for communicating science in languages with fewer speakers (for which large language models might not yet exist), generative AI may be a dead end for now. The lack of cultural nuance, ignorance of current events and inability to contextualize with human expertise – all of these make generative AI a non-starter for certain science communication use cases.

S.S.H.: For science communicators, the ability to convey information in ways that resonate with specific cultural and regional audiences is crucial for enhancing the collective understanding of scientific issues. At the point of writing, GPT-4 (for example) has not fully comprehended the cultural intricacies of its content generation and is unable to account for minority languages, dialects,

colloquialisms and regional idioms. Studies have shown that English GPT-4 closely reflects the cultural values of the USA when it is used to answer contextual questions in other countries¹⁶. This suggests the possibility of cultural incompatibility, particularly with non-Western countries. Consequently, human intervention is necessary because the cultural compatibility of generative AI tools is still questionable.

What do generative AI tools mean for creativity in science communication?

A.A.: During a similar, past technological upheaval – the invention of photography – it was feared that painting, a creative art that imperfectly captured the world yet required technical mastery, would be destroyed.

Now, using generative AI tools such as [Midjourney](#), one can achieve hypersurreal capture of any idea – from prompt to picture – with zero technical aptitude.

There may be a silver lining for science communication in this unprecedented shift. Generative AI tools are systems of statistical mimicry and vast combinatorial capacity. This opens a treasure trove for scientific storytelling and outputs, and also for creativity – a huge remix dashboard in which only the cleverness of a prompt sets the limit. Illustrations for stories can be generated with just a one-line idea; media such as cartoons or audio become accessible, with storyboards or voiceovers generated in seconds; and the creation and iteration of pilot ideas make it easier than ever to overcome writer's block. There are drawbacks

to this creative freewheeling, though; creative collaboration with AI reduces people's sense of ownership¹⁷. The speed and exploratory power of generative AI tools must be balanced by human goals and agency, both for the creator and consumer of science communication.

Amanda Alvarez¹✉, **Aylin Caliskan**²✉,
M. J. Crockett^{3,4}✉, **Shirley S. Ho**⁵✉,
Lisa Messeri⁶✉ & **Jevin West**^{2,7}✉

¹Finnish Center for Artificial Intelligence, Aalto University, Espoo, Finland. ²The Information School, University of Washington, Seattle, WA, USA. ³Department of Psychology, Princeton University, Princeton, NJ, USA.

⁴University Center for Human Values, Princeton University, Princeton, NJ, USA. ⁵Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore.

⁶Department of Anthropology, Yale University, New Haven, CT, USA. ⁷Center for an Informed Public, University of Washington, Seattle, WA, USA.

✉e-mail: amanda.alvarez@aalto.fi; aylin@uw.edu; mj.crockett@princeton.edu; tsyho@ntu.edu.sg; lisa.messeri@yale.edu; jevinw@uw.edu

Published online: 04 March 2024

References

1. Wang, H. et al. *Nature* **620**, 47–60 (2023).
2. Wagner, G., Lukyanenko, R. & Paré, G. *J. Inf. Technol.* **37**, 209–226 (2022).
3. Yang, Y., Youyou, W. & Uzzi, B. *Proc. Natl Acad. Sci. USA* **117**, 10762–10768 (2020).

Published online: 04 March 2024

References

1. Wang, H. et al. *Nature* **620**, 47–60 (2023).
2. Wagner, G., Lukyanenko, R. & Paré, G. *J. Inf. Technol.* **37**, 209–226 (2022).
3. Yang, Y., Youyou, W. & Uzzi, B. *Proc. Natl Acad. Sci. USA* **117**, 10762–10768 (2020).

4. Messeri, L. & Crockett, M. J. Artificial intelligence and illusions of understanding in scientific research. *Nature*, <https://doi.org/10.1038/s41586-024-07146-0> (2024).
5. Ghosh, S. & Caliskan, A. ChatGPT perpetuates gender bias in machine translation and ignores non-gendered pronouns: findings across Bengali and five other low-resource languages. In *AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society* (eds Rossi, F. et al.) 901–912 (AAAI/ACM AIES 2023).
6. Longino, H. E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Princeton Univ. Press, 1990).
7. Harding, S. *Centen. Rev.* **36**, 437–470 (1992).
8. Nakadai, R., Nakawake, Y. & Shibasaki, S. *Nat. Hum. Behav.* **7**, 1804–1805 (2023).
9. Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big? In *Proc. 2021 ACM Conf. on Fairness, Accountability, and Transparency*, 610–623 (ACM, 2021).
10. Crockett, M. J. & Messeri, L. Preprint at OSF, <https://doi.org/10.31234/osf.io/4zdx9> (2023).
11. Conroy, G. *Nature* **622**, 234–236 (2023).
12. Ou, M. & Ho, S. S. *Public Underst. Sci.* **33**, 241–259 (2023).
13. Ho, S. S., Scheufele, D. A. & Corley, E. A. *Sci. Commun.* **33**, 167–200 (2011).
14. Ho, S. S. et al. *Environ. Commun.* **13**, 457–471 (2019).
15. Shin, D., Rasul, A. & Fotiadis, A. *Internet Res.* **32**, 1214–1234 (2022).
16. Cao, Y. et al. (2023). Assessing cross-cultural alignment between ChatGPT and human societies: an empirical study. In *Proc. First Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, 53–67 (ACL, 2023).
17. Draxler, F. et al. (2023). The AI ghostwriter effect: when users do not perceive ownership of AI-generated text but self-declare as authors. In *ACM Trans. Computer–Human Interaction*, vol. 31 (eds Höök, E. & Hornbæk, K.) 25 (ACM, 2023).

Acknowledgements

S.S.H. is part of the research programme DesCartes, which is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

Competing interests

J.W. is on the board of Consensus.app. The other authors declare no competing interests.