HOW GENERATIVE AI COULD DISRUPT SCIENTIFIC PUBLISHING

A world of AI-assisted writing and reviewing might transform the nature of the scientific paper. **By Gemma Conroy**

When radiologist Domenico Mastrodicasa finds himself stuck while writing a research paper, he turns to ChatGPT, the chatbot that produces fluent responses to almost any query in seconds. "I use it as a sounding board," says Mastrodicasa, who is based at the University of Washington School of Medicine in Seattle. "I can produce a publication-ready manuscript much faster."

Mastrodicasa is one of many researchers experimenting with generative artificial-intelligence (AI) tools to write text or code. He pays for ChatGPT Plus, the subscription version of the bot based on the large language model (LLM) GPT-4, and uses it a few times a week. He finds it particularly useful for suggesting clearer ways to convey his ideas. Although a *Nature* survey suggests that scientists who use LLMs regularly are still in the minority, many expect that generative AI tools will become regular assistants for writing manuscripts, peer-review reports and grant applications.

Those are just some of the ways in which AI could transform scientific communication and publishing. Science publishers are already experimenting with generative AI in scientific search tools and for editing and quickly summarizing papers. Many researchers think that non-native English speakers could benefit most from these tools. Some see generative AI as a way for scientists to rethink how they interrogate and summarize experimental results altogether – they could use LLMs to do much of this work, meaning less time writing papers and more time doing experiments.

"It's never really the goal of anybody to write papers – it's to do science," says Michael Eisen, a computational biologist at the University of California, Berkeley, who is also editor-in-chief of the journal *eLife*. He predicts that generative AI tools could even fundamentally transform the nature of the scientific paper.

But the spectre of inaccuracies and falsehoods threatens this vision. LLMs are merely engines for generating stylistically plausible output that fits the patterns of their inputs, rather than for producing accurate information. Publishers worry that a rise in their use might lead to greater numbers of poor-quality or error-strewn manuscripts – and possibly a flood of AI-assisted fakes.

"Anything disruptive like this can be quite worrying," says Laura Feetham, who oversees peer review for IOP Publishing in Bristol, UK, which publishes physical-sciences journals.

A flood of fakes?

Science publishers and others have identified a range of concerns about the potential impacts of generative AI. The accessibility of generative AI tools could make it easier to whip up poor-quality papers and, at worst, compromise research integrity, says Daniel Hook, chief executive of Digital Science, a research-analytics firm in London. "Publishers are quite right to be scared," says Hook. (Digital Science is part of Holtzbrinck Publishing Group, the majority shareholder in *Nature*'s publisher, Springer Nature; *Nature*'s news team is editorially independent.)

In some cases, researchers have already admitted using ChatGPT to help write papers without disclosing that fact. They were caught because they forgot to remove telltale signs of its use, such as fake references or the software's preprogrammed response that it is an AI language model.

Ideally, publishers would be able to detect LLM-generated text. In practice, AI-detection tools have so far proved unable to pick out such text reliably while avoiding flagging human-written prose as the product of an AI.

Although developers of commercial LLMs are working on watermarking LLM-generated output to make it identifiable, no firm has yet rolled this out for text. Any watermarks could also be removed, says Sandra Wachter, a legal

> "I think it's inevitable that this will be part of our toolkit."

scholar at the University of Oxford, UK, who focuses on the ethical and legal implications of emerging technologies. She hopes that lawmakers worldwide will insist on disclosure or watermarks for LLMs, and will make it illegal to remove watermarking.

Publishers are approaching the issue either by banning the use of LLMs altogether (as *Science*'s publisher, the American Association for the Advancement of Science, has done), or, in most cases, insisting on transparency (the policy at *Nature* and many other journals). A study examining 100 publishers and journals found that, as of May, 17% of publishers and 70% of journals had released guidelines on how generative AI could be used, although they varied on how the tools could be applied, says Giovanni Cacciamani, a urologist at the University of Southern California in Los Angeles, who co-authored the work, which has not yet been peer reviewed[1]. He and his colleagues are working with scientists and journal editors to develop a uniform set of guidelines to help researchers to report their use of LLMs.

Many editors are concerned that generative AI could be used to more easily produce fake but convincing articles. Companies that create and sell manuscripts or authorship positions to researchers who want to boost their publishing output, known as paper mills, could stand to profit. A spokesperson for *Science* told *Nature* that LLMs such as ChatGPT could exacerbate the paper-mill problem.

One response to these concerns might be for some journals to bolster their approaches to verify that authors are genuine and have done the research they are submitting. "It's going to be important for journals to understand whether or not somebody actually did the thing they are claiming," says Wachter.

At the publisher EMBO Press in Heidelberg, Germany, authors must use only verifiable institutional e-mail addresses for submissions, and editorial staff meet with authors and referees in video calls, says Bernd Pulverer, head of scientific publications there. But he adds that research institutions and funders also need to monitor the output of their staff and grant recipients more closely. "This is not something that can be delegated entirely to journals," he says.

Equity and inequity

When *Nature* surveyed researchers on what they thought the biggest benefits of generative AI might be for science, the most popular answer was that it would help researchers who do not have English as their first language (see *Nature* **621**, 672–675; 2023). "The use of AI tools could improve equity in science," says Tatsuya Amano, a conservation scientist at the University of Queensland in Brisbane, Australia. Amano and his colleagues surveyed more than 900 environmental scientists who had authored at least one paper in English. Among early-career researchers, non-native English speakers said their papers were rejected owing to writing issues more than twice as often as native English speakers did, who also spent less time writing their submissions[2]. ChatGPT and similar tools could be a "huge help" for these researchers, says Amano.

Amano, whose first language is Japanese, has been experimenting with ChatGPT and says the process is similar to working with a native English-speaking colleague, although the tool's suggestions sometimes fall short. He co-authored an editorial in *Science* in March following that journal's ban on generative AI tools, arguing that they could make scientific publishing more equitable as long as authors disclose their use, such as by including the original manuscript alongside an AI-edited version[3].

LLMs are far from the first AI-assisted software that can polish writing. But generative AI is simply much more flexible, says Irene Li, an AI researcher at the University of Tokyo. She previously used Grammarly – an AI-driven grammar and spelling checker – to improve her written English, but has since switched to ChatGPT because it's more versatile and offers better value in the long run; instead of paying for multiple tools, she can subscribe to just one that does it all. "It saves a lot of time," she says.

However, the way in which LLMs are developed might exacerbate inequities, says Chhavi Chauhan, an AI ethicist who is also director of

scientific outreach at the American Society for Investigative Pathology in Rockville, Maryland. Chauhan worries that some free LLMs might become expensive in the future to cover the costs of developing and running them, and that if publishers use AI-driven detection tools, they are more likely to erroneously flag text written by non-native English speakers as AI. A study in July found this does happen with the current generation of GPT detectors[4]. "We are completely missing the inequities these generative AI models are going to create," she says.

Peer-review challenges

LLMs could be a boon for peer reviewers, too. Since using ChatGPT Plus as an assistant, Mastrodicasa says he's been able to accept more review requests, using the LLM to polish his comments, although he doesn't upload manuscripts or any information from them into the online tool. "When I already have a draft, I can refine it in a few hours rather than a few days," he says. "I think it's inevitable that this will be part of our toolkit." Christoph Steinbeck, a chemistry informatics researcher at the Friedrich Schiller University in Jena, Germany, has found ChatGPT Plus handy for creating quick summaries for preprints he's reviewing. He notes that preprints are already online, and so confidentiality is not an issue.

One key concern is that researchers could rely on ChatGPT to whip up reviews with little thought, although the naive act of asking an LLM directly to review a manuscript is likely to produce little of value beyond summaries and copy-editing suggestions, says Mohammad Hosseini, who studies research ethics and integrity at Northwestern University's Galter Health Sciences Library and Learning Center in Chicago, Illinois.

Most of the early worries over LLMs in peer review have been about confidentiality. Several publishers — including Elsevier, Taylor & Francis and IOP Publishing — have barred researchers from uploading manuscripts and sections of text to generative AI platforms to produce peer-review reports, over fears that the work might be fed back into an LLM's training data set, which would breach contractual terms to keep work confidential. In June, the US National Institutes of Health banned the use of ChatGPT and other generative AI tools to produce peer reviews of grants, owing to confidentiality concerns. Two weeks later, the Australian Research Council prohibited the use of generative AI during grant review for the same reason, after a number of reviews that seemed to be written by ChatGPT appeared online.

One way to get around the confidentiality hurdle is to use privately hosted LLMs. With these, one can be confident that data are not fed back to the firms that host LLMs in the cloud. Arizona State University in Tempe is experimenting with privately hosted LLMs

based on open-source models, such as Llama 2 and Falcon. "It's a solvable problem," says Neal Woodbury, chief science and technology officer at the university's Knowledge Enterprise, who advises university leaders on research initiatives.

Feetham says that if it was clearer how LLMs store, protect and use the data that are put into them, then the tools could conceivably be integrated into the reviewing systems that publishers already use. "There are real opportunities there if these tools are used properly." Publishers have been using machine-learning and natural-language-processing AI tools to assist with peer review for more than half a decade, and generative AI could augment the capabilities of this software. A spokesperson for the publisher Wiley says the company is experimenting with generative AI to help screen manuscripts, select reviewers and verify the identity of authors.

Ethical concerns

Some researchers, however, argue that LLMs are too ethically murky to include in the scientific publishing process. A main concern lies

"Anything disruptive like this can be quite worrying."

in the way LLMs work: by trawling Internet content without concern for bias, consent or copyright, says Iris van Rooij, a cognitive scientist at Radboud University in Nijmegen, the Netherlands. She adds that generative AI is "automated plagiarism by design", because users have no idea where such tools source their information from. If researchers were more aware of this problem, they wouldn't want to use generative AI tools, she argues.

Some news organizations have blocked ChatGPT's bot from trawling their sites, and media reports suggest that some firms are contemplating lawsuits. Although scientific publishers haven't gone that far in public, Wiley told *Nature* that it was "closely monitoring industry reports and litigation claiming that generative AI models are harvesting protected material for training purposes while disregarding any existing restrictions on that information". The publisher also noted that it had called for greater regulatory oversight, including transparency and audit obligations for providers of LLMs.

Hosseini, who is also an assistant editor for the journal *Accountability in Research*, which is published by Taylor & Francis, suggests that training LLMs on scientific literature in specific disciplines could be one way to improve both the accuracy and relevance of their output to scientists — although no publishers contacted by *Nature* said they were doing this.

If scholars start to rely on LLMs, another concern is that their expression skills might atrophy, says Gemma Derrick, who studies research policy and culture at the University of Bristol, UK. Early-career researchers could miss out on developing the skills to conduct fair and balanced reviews, she says.

Transformational change

More broadly, generative AI tools have the potential to change how research is published and disseminated, says Patrick Mineault, a senior machine-learning scientist at Mila — Quebec AI Institute in Montreal, Canada. That could mean that research will be published in a way that can be easily read by machines rather than humans. "There will be all these new forms of publication," says Mineault.

In the age of LLMs, Eisen pictures a future in which findings are published in an interactive, "paper on demand" format rather than as a static, one-size-fits-all product. In this model, users could use a generative AI tool to ask queries about the experiments, data and analyses, which would allow them to drill into the aspects of a study that are most relevant to them. It would also allow users to access a description of the results that is tailored to their needs. "I think it's only a matter of time before we stop using single narratives as the interface between people and the results of scientific studies," says Eisen.

Companies such as scite and Elicit have already launched search tools that use LLMs to provide researchers with natural-language answers to queries; in August, Elsevier launched a pilot version of its own tool, Scopus AI, to give quick summaries of research topics. Generally, these tools use LLMs to rephrase results that come back from conventional search queries.

Mineault adds that generative AI tools could change how researchers conduct meta-analyses and reviews — although only if the tools' tendency to make up information and references can be addressed adequately. The largest human-generated review that Mineault has seen included around 1,600 papers, but working with generative AI could take it much further. "That's a very tiny proportion of the whole scientific literature," he says. "The question is, how much stuff is in the scientific literature right now that could be exploited?"

Gemma Conroy is a reporter for *Nature* in Sydney, Australia.

1. Ganjavi, C. *et al.* Preprint at https://arxiv.org/abs/2307.11918 (2023).
2. Amano, T. *et al.* PLoS Biol. **21**, e3002184 (2023).
3. Berdejo-Espinola, V. & Amano, T. Science **379**, 991 (2023).
4. Liang, W. *et al.* Patterns **4**, 100779 (2023).