

How language-generating AIs could transform science

Machine-learning algorithms that generate fluent language from vast amounts of text could change how science is done — but not necessarily for the better, says Shobita Parthasarathy, a specialist in the governance of emerging technologies at the University of Michigan in Ann Arbor. In a report published on 27 April (see go.nature.com/381jjiv), Parthasarathy and other researchers try to anticipate societal impacts of emerging artificial-intelligence (AI) technologies called large language models (LLMs). These can churn out astonishingly convincing prose, translate between languages, answer questions and even produce code. Parthasarathy says that although LLMs could strengthen efforts to understand complex research, they could also deepen public scepticism of science.

How might LLMs help or hinder science?

I had originally thought that LLMs could have democratizing and empowering impacts. When it comes to science, they could empower people to quickly pull insights out of information: by querying disease symptoms, for example, or generating summaries of technical topics.

But the algorithmic summaries could make errors, include outdated information or remove nuance and uncertainty, without users appreciating this. If anyone can use LLMs to make complex research comprehensible, but they risk getting a simplified, idealized view of science that's at odds with the messy reality, that could threaten professionalism and authority. It might also exacerbate problems of public trust in science. And people's interactions with these tools will be very individualized, with each user getting their own generated information.

Isn't the possibility that LLMs might draw on outdated or unreliable research a huge problem?

Yes. But that doesn't mean people won't use LLMs. They're enticing, and they will have a veneer of objectivity associated with their fluent output and their portrayal as exciting new technologies. The fact that they have limits — that they might be built on partial or historical data sets — might not be recognized by the average user.

It's easy for scientists to assert that they



Shobita Parthasarathy says that the use of large language models should be regulated.

are smart and realize that LLMs are useful. Still, these kinds of tool could narrow their field of vision, and it might be hard to recognize when an LLM gets something wrong.

LLMs could be useful in digital humanities, for instance: to summarize what a historical text says about a particular topic. But these models' processes are opaque, and they don't provide sources alongside their outputs, so researchers will need to think carefully about how they're going to use them.

Who might create these models for science?

My guess is that large scientific publishers are going to be in the best position to develop science-specific LLMs able to crawl over the proprietary full text of their papers. They could also look to automate aspects of peer review, such as querying scientific texts to find out who should be consulted as a reviewer. LLMs might also be used to try to pick out particularly innovative results in manuscripts, and perhaps even to help evaluate these results.

Could LLMs be used to make realistic but fake papers?

Yes — some people will use LLMs to generate fake or near-fake papers, if it is easy and they think that it will help their career. Still, that doesn't mean that most scientists, who do

want to be part of scientific communities, won't be able to agree on regulations and norms for using LLMs.

How should the use of LLMs be regulated?

It's fascinating to me that hardly any AI tools have been put through systematic regulations. That's true for LLMs, too: their methods are opaque and vary by developer. In our report, we make recommendations for government bodies to step in with general regulation.

Specifically for LLMs' possible use in science, transparency is crucial. Those developing LLMs should explain what texts have been used and the logic of the algorithms involved — and should be clear about whether computer software has been used to generate an output.

And scientists should be wary of journals or funders relying on LLMs for finding peer reviewers or (conceivably) extending this process to other aspects of review, such as evaluating manuscripts or grants. Because LLMs veer towards past data, they are likely to be too conservative in their recommendations.

Interview by Richard Van Noorden

This interview has been edited for length and clarity.