World view

By Arthur Spirling

Open generative AI models are a way forward for science

Researchers should stop using proprietary large language models and develop transparent ones to ensure reproducibility.

Every day, it seems, a new large language model (LLM) is announced with breathless commentary — from both its creators and academics — on its extraordinary abilities to respond to human prompts. It can fix code! It can write a reference letter! It can summarize an article!

From my perspective as a political and data scientist who is using and teaching about such models, scholars should be wary. The most widely touted LLMs are proprietary and closed: run by companies that do not disclose their underlying model for independent inspection or verification, so researchers and the public don't know on which documents the model has been trained.

The rush to involve such artificial-intelligence (AI) models in research is a problem. Their use threatens hard-won progress on research ethics and the reproducibility of results.

Instead, researchers need to collaborate to develop open-source LLMs that are transparent and not dependent on a corporation's favours.

It's true that proprietary models are convenient and can be used out of the box. But it is imperative to invest in open-source LLMs, both by helping to build them and by using them for research. I'm optimistic that they will be adopted widely, just as open-source statistical software has been. Proprietary statistical programs were popular initially, but now most of my methodology community uses open-source platforms such as R or Python.

One open-source LLM, BLOOM, was released last July. BLOOM was built by New York City-based AI company Hugging Face and more than 1,000 volunteer researchers, and partially funded by the French government. Other efforts to build open-source LLMs are under way. Such projects are great, but I think we need even more collaboration and pooling of international resources and expertise. Open-source LLMs are generally not as well funded as the big corporate efforts. Also, they need to run to stand still: this field is moving so fast that versions of LLMs are becoming obsolete within weeks or months. The more academics who join these efforts, the better.

Using open-source LLMs is essential for reproducibility. Proprietors of closed LLMs can alter their product or its training data — which can change its outputs — at any time.

For example, a research group might publish a paper testing whether phrasings suggested by a proprietary LLM can help clinicians to communicate more effectively with patients. If another group tries to replicate that study, who knows whether the model's underlying training data will be the same, or even whether the technology will still be supported? GPT-3, released last November by OpenAI in San Francisco, California, has already been supplanted by GPT-4, and presumably supporting the older LLM will soon no longer be the firm's main priority.

> "With open-source large language models, researchers can look at the guts of the model to see how it works, customize its code and flag errors."

By contrast, with open-source LLMs, researchers can look at the guts of the model to see how it works, customize its code and flag errors. These details include the model's tunable parameters and the data on which it was trained. Engagement and policing by the community help to make such models robust in the long term.

The use of proprietary LLMs in scientific studies also has troubling implications for research ethics. The texts used to train these models are unknown: they might include direct messages between users on social-media platforms or content written by children legally unable to consent to sharing their data. Although the people producing the public text might have agreed to a platform's terms of service, this is perhaps not the standard of informed consent that researchers would like to see.

In my view, scientists should move away from using these models in their own work where possible. We should switch to open LLMs and help others to distribute them. Moreover, I think academics, especially those with a large social-media following, shouldn't be pushing others to use proprietary models. If prices were to shoot up, or companies fail, researchers might regret having promoted technologies that leave colleagues trapped in expensive contracts.

Researchers can currently turn to open LLMs produced by private organizations, such as LLaMA, developed by Facebook's parent company Meta in Menlo Park, California. LLaMA was originally released on a case-by-case basis to researchers, but the full model was subsequently leaked online. My colleagues and I are working with Meta's open LLM OPT-175B, for instance. Both LLaMA and OPT-175B are free to use. The downside in the long run is that this leaves science relying on corporations' benevolence — an unstable situation.

There should be academic codes of conduct for working with LLMs, as well as regulation. But these will take time and, in my experience as a political scientist, I expect that such regulations will initially be clumsy and slow to take effect.

In the meantime, massive collaborative projects urgently need support to produce open-source models for research — like CERN, the international organization for particle physics, but for LLMs. Governments should increase funding through grants. The field is moving at lightning speed and needs to start coordinating national and international efforts now. The scientific community is best placed to assess the risks of the resulting models, and might need to be cautious about releasing them to the public. But it is clear that the open environment is the right one.

Arthur Spirling is a professor of politics and data science at New York University.
e-mail: as9934@nyu.edu

ARTHUR SPIRLING