

# New language-learning algorithms risk reinforcing inequalities, social fragmentation

April 28 2022, by Daniel Rivkin

---



Credit: Pixabay/CC0 Public Domain

The use of large language models could transform many facets of modern life, including how policymakers assess public sentiment about pending legislation, how patients evaluate their medical care and how

scientists could translate research findings across languages.

Yet, new research from the University of Michigan finds that while there's great potential for these machine learning algorithms to benefit society, they likely could reinforce inequalities, tax the environment and place still more power in the hands of tech giants.

Large language models, or LLMs, can recognize, summarize, translate, predict and generate [human languages](#) on the basis of very large text-based datasets, and are likely to provide the most convincing computer-generated imitation of human language yet.

A report by the Technology Assessment Project at the Science, Technology, and Public Policy (STPP) program at the Gerald R. Ford School of Public Policy raises concerns about the many ways that LLMs can cause profoundly negative outcomes.

The report, "What's in the Chatterbox? Large Language Models, Why They Matter, and What We Should Do About Them," anticipates the transformative social change they could produce:

- Because of the concentrated development landscape and the nature of LLM datasets, the new technologies will not represent marginalized communities adequately. They are likely to systematically minimize and misrepresent these voices while amplifying the perspectives of the already powerful.
- LLM processing occurs in physical data centers, which require massive amounts of natural resources. Data center construction is already disproportionately harming marginalized populations.
- LLMs will accelerate [tech companies'](#) thirst for data, become quickly integrated into existing information infrastructure, reorganize labor and expertise, reinforce inequality and increase social fragmentation.

"Our analysis shows that LLMs could empower communities and democratize knowledge, but right now they are unlikely to achieve this potential. The harms can be mitigated, but not without new rules and regulations about how these technologies are created and used," said STPP director Shobita Parthasarathy, professor of [public policy](#).

The report uses the analogical case study method to analyze LLM development and adoption, by examining the history of similar past technologies—in terms of form, function and impacts—to anticipate the implications of emerging technologies. STPP pioneered this method in previous reports on facial recognition technologies in K-12 schools and vaccine hesitancy.

"Technologies can be implemented widely and then the negative consequences can take years to correct. LLMs present many of the same equity, environmental and access issues we have seen in previous cases," said Johanna Okerlund, STPP postdoctoral fellow and report co-author.

LLMs are much larger than their artificial intelligence predecessors, both in terms of the massive amounts of data developers use to train them and the millions of complex word patterns and associations the models contain. They are more advanced than previous natural language processing efforts because they can complete many types of tasks without being specifically trained for each, which makes any single LLM widely applicable.

Numerous factors create the circumstances for built-in inequity, according to the report.

"LLMs require enormous resources in terms of finances, infrastructure, personnel and computational resources including 360,000 gallons of water a day and immense electricity, infrastructure and rare earth material usage," the report says.

Only a handful of tech companies can afford to build them, and their construction is likely to disproportionately burden already marginalized communities. The authors also say they worry "because LLM design is likely to distort or devalue the needs of marginalized communities ... LLMs might actually alienate them further from social institutions."

Researchers also note the vast majority of models are based on texts in English, and, to a lesser extent, Chinese.

"This means that LLMs are unlikely to achieve their translation goals (even to and from English and Chinese) and will be less useful for those who are not English or Chinese dominant," the report says.

One example of the analogical case study method's utility is to examine how [racial bias](#) is already embedded in many medical devices including the spirometer, which is used to measure lung function: "The technology considers race in its assessment of 'normal' lung function, falsely assuming that Black people naturally have lower lung function than their white counterparts, and making it more difficult for them to access treatment."

"We expect similar scenarios in other domains including criminal justice, housing and education, where biases and discrimination enshrined in historical texts are likely to generate advice that perpetuates inequities in resource allocation," the report says.

"LLMs' thirst for data will jeopardize privacy, and customary methods for establishing informed consent will no longer work.

"Because they collect enormous amounts of data, LLMs will likely be able to triangulate bits of disconnected information about individuals including mental health status or political opinions to develop a full, personalized picture of actual people, their families or communities. In a



world with LLMs, the customary method for ethical data collection—individual informed consent—no longer makes sense" and can cross to unethical methods of data collection in order to diversify the [data sets](#).

LLMs will affect many sectors, but the report dives deeply into one to provide an example: How they will influence [scientific research](#) and practice. The authors suggest that academic publishers, which own most research publications, will construct their own LLMs and use them to increase their monopoly power.

Meanwhile, researchers will need to develop standard protocols on how to scrutinize insights generated by LLMs and how to cite output so others can replicate the results. Scientific inquiry will likely shift to finding patterns in big data rather than establishing causal relationships. And scientific evaluation systems relying on LLMs will probably not be able to identify truly novel work, a task that is already quite difficult for human beings.

Given these likely outcomes, the authors suspect scientists will come to distrust LLMs.

The report concludes with [policy recommendations](#), which include:

- U.S. government regulation of LLMs, including a clear definition of what constitutes an LLM, evaluation and approval protocols based on content and algorithms, and security, oversight and complaint mechanisms.
- Regulation of apps that use LLMs.
- National or international standards that look at data set diversity, performance, transparency, accuracy, security and bias, as well as copyright protection of LLM-generated inventions and artistic works.

- Methods of ensuring security and personal privacy when deploying LLMs particularly among vulnerable populations.
- Full-time government advisers in the social and equity dimensions of technology, including a "Chief Human Rights in Tech Officer."
- Environmental assessments of new data centers that evaluate the impacts on local utility prices, local marginalized communities, human rights in minerals mining and climate change.
- Evaluate the health, safety and psychological risks that LLMs and other forms of artificial intelligence create for workers, e.g., reorienting them towards more complex and often unsafe tasks, and developing a response to the job consolidation that LLMs, and automation more generally, are likely to create.
- A call for the National Science Foundation to substantially increase its funding for LLM development, with a focus on the equity, social and environmental impacts of LLMs.

The report also outlines specific recommendations for the scientific community and a Developer's Code of Conduct.

"Both LLM and app developers must recognize their public responsibilities and try to maximize the benefits of these technologies while minimizing the risks," the authors wrote.

Provided by University of Michigan

Citation: New language-learning algorithms risk reinforcing inequalities, social fragmentation (2022, April 28) retrieved 23 April 2024 from <https://techxplore.com/news/2022-04-language-learning-algorithms-inequalities-social-fragmentation.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.