Environmental Science & Technology

pubs.acs.org/est

Viewpoint

Risks and Benefits of Large Language Models for the Environment

Matthias C. Rillig,* Marlene Ågerstrand, Mohan Bi, Kenneth A. Gould, and Uli Sauerland

Cite This: *Environ. Sci. Technol.* 2023, 57, 3464–3466

Read Online

ACCESS | Metrics & More | Article Recommendations

KEYWORDS: *large language models, ChatGPT, environmental education, environmental impact, digital divide, environmental literacy, artificial intelligence*

Large language models (LLMs) are artificial intelligence (AI) models with complex architecture and a large number of parameters that have been trained on very large amounts of text (billions of words). These models, arising from the field of natural language processing, can generate natural, human-like writing and have been designed to react to user input, enabling conversations and customized output according to prompts. The release of the chatbot ChatGPT by OpenAI[1] in late 2022 has rapidly spread this technology to a wide range of users (GPT means generative pretrained transformer and denotes the type of large language model used in this chatbot). Other companies are offering their own apps featuring different kinds of LLMs, and this technology is also being rapidly integrated into existing apps and online tools. Because LLMs will likely become extremely common, the potentially transformative nature of these models has already sparked a lively debate about their use. This discussion focuses on academic integrity and the future of research and teaching, the meaning of authorship, potential consequences for the general workforce, and unresolved copyright issues.[2] However, the debate has so far largely missed the potential implications of current and future LLM tools for the environment.[3] We see the possibility of direct and indirect environmental impacts and effects, and opportunities and risks for researchers in the environmental sciences (Figure 1).

The first point to consider is the positive or negative direct environmental impact (Figure 1). Other potentially transformative technological innovations, such as the metaverse,[4] likely will have direct consequences on the environment via increased energy use and thus resource consumption and production of carbon dioxide. This clearly is also a concern for LLMs,[5] with both the training of LLMs and inference having large energy demands, prompting an early call for algorithmic efficiency.[3] The carbon footprint will depend on the energy use and the carbon intensity of the energy source being used. In addition to carbon dioxide emissions, the computing facilities may also have other environmental impacts such as water use and soil pollution or sealing, which could have broader implications for environmental quality. Conversely, can the use of text-based chats in the future partially replace video conferences or travel to in-person meetings that might consume more resources by comparison? It is unclear if this will be the case, given that human verification and expertise will likely remain indispensable.[2]

There are likely also indirect consequences of increased LLM use, which are potentially more important (Figure 1). The first issue is the level of artificial expertise. LLM output comes with a certain degree of simulated authority, given the extensive amount of information with which LLMs have been trained and the polished language in which output is written. Therefore, such output can easily be confused with expert opinions, even though LLMs will continue to have limited ability to judge the reliability and relevance of information, in part because LLMs have not achieved natural language *understanding*. Thus, false output is created, as anybody who has played with these apps on topics of their own expertise will have noticed. More worrying is the potential to inadvertently or purposefully introduce bias at three points: the training data (the input to the model), the algorithm (how sources are used), and the form of output (e.g., disclaimers, statements of uncertainty, and references). At each of these points, special interest groups and networks could exploit the ability of LLMs to generate text with unprecedented efficiency, thus offering misinformation under the guise of "artificial intelligence", and flooding public spaces with it. We think this is the biggest concern of the more widespread use of LLMs for environmentally relevant topics. But even without ill intent, existing biases on complex environmental topics, including environmental racism,[3] climate change and global environmental change, biodiversity loss, or pollution, could be perpetuated and multiplied by the training data the LLM uses. On the contrary, writing informative content about environmental issues by actors interested in environmental education could be made more efficient through LLMs. For example, materials for use in environmental education could be more easily adapted

Published: February 23, 2023

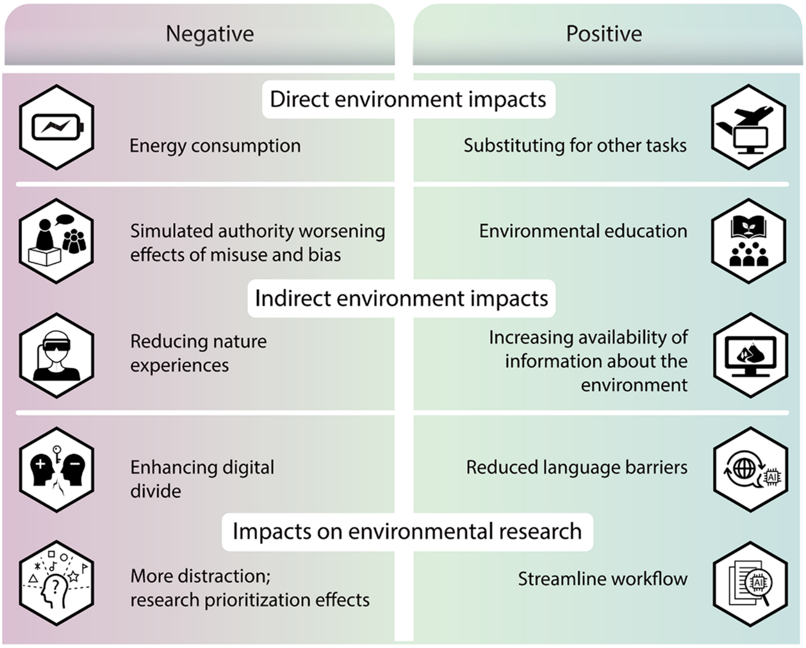

Figure 1. Large language models come with risks and opportunities for the environment. Increased use of large language models could affect the environment positively or negatively, with possible direct and indirect effects on the environment and on the way environmental research is conducted.

for different target groups, such as different ages or educational levels.

Differential access to LLM-based apps could worsen or improve digital divide effects within and among societies. The availability of these tools could further favor people with already excellent access to environmental information. On the positive side, LLMs could broaden the participation of people in the environmental debate, especially because the LLM services are being offered in various languages.

It is also possible that increasingly relying on technology-guided interactions could contribute to reduced nature experiences,[6] thus reducing value attributed to the environment. More screen time spent on fascinating tools could translate to less time spent in nature, with potential consequences for how people will appreciate biodiversity and ecosystems. Conversely, the public may also benefit from unprecedented, up-to-date, accessible, and personalized information and educational opportunities on environmental topics. This increasingly available "information at your fingertips" could lead to greater curiosity about environmental topics, thus enhancing environmental literacy.

There are certainly many benefits of LLMs for research in the environmental sciences. The efficient use of these tools will likely streamline the workflow of environmental scientists and potentially improve the quality of writing, thus freeing up researchers' time for other tasks, i.e., designing and analyzing experiments and developing innovative ideas. By offering a tool to sharpen their English language scientific writing, LLMs might increase representation of researchers from non-English speaking countries in the environmental sciences. This all could potentially accelerate scientific progress on important environmental topics. Conversely, these tools might prove to be a distraction, and researcher time may increasingly be consumed by dealing with adverse effects of LLM use, for example, in university education and training (e.g., misuse and cheating).[2] If LLMs are increasingly used to summarize information and to support funding decisions in the future, this could render human expertise and insight less important, potentially leading to adverse effects on research prioritization in the environmental sciences.

Clearly, there are benefits and risks for the environment in the increased use of LLMs, and it is important to start this discussion early to avert possible harm. To harness the opportunities, we ought to protect such powerful tools from the undue influence of individual groups. To ensure that LLMs are constructed to yield unbiased information, governments, intergovernmental bodies, and large organizations (such as OECD) need to devise policy tools that could include oversight committees, guidelines, or regulations that lead to disclosure of data sources (for training of models) and funding. Another important goal is to increase literacy in the use of these tools, and this is where environmental scientists can act now, by including LLMs in their university courses and by clarifying in their laboratories how LLMs can and should be used in the environmental sciences.

AUTHOR INFORMATION

Corresponding Author

Matthias C. Rillig − *Freie Universität Berlin, Institute for Biology, 14195 Berlin, Germany; Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), 14195 Berlin, Germany;* orcid.org/0000-0003-3541-7853; Email: matthias.rillig@fu-berlin.de

Authors

Marlene Ågerstrand − *Department of Environmental Science, Stockholm University, SE-106 91 Stockholm, Sweden*

Mohan Bi − *Freie Universität Berlin, Institute for Biology, 14195 Berlin, Germany; Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), 14195 Berlin, Germany*

Kenneth A. Gould – *Department of Sociology, Brooklyn College, City University of New York, Brooklyn, New York 11210, United States*
Uli Sauerland – *Leibniz-Centre General Linguistics (ZAS), 10117 Berlin, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.est.3c01106

Author Contributions

M.C.R. wrote the first draft of the paper. All co-authors contributed ideas and edited the text. M.B. designed the figure.

Notes

The authors declare no competing financial interest.

Biography

Matthias C. Rillig studied biology in Germany and Scotland and obtained a Ph.D. in California, USA. After 9 years of being on the faculty of the University of Montana, he joined Freie Universität Berlin, where he is now a professor of ecology. He is director of the Berlin-Brandenburg Institute of Advanced Biodiversity Research. Matthias won an Advanced Grant of the European Research Council, is a fellow of the Ecological Society of America, and is a member of the German National Academy of Sciences, Leopoldina, and Academia Europaea. His lab focuses on soil ecology and human-caused effects on soils and their biodiversity.

REFERENCES

(1) *ChatGPT: Optimizing Language Models for Dialogue*; OpenAI. https://openai.com/blog/chatgpt/ (accessed 2023-01-16).
(2) van Dis, E. A. M.; Bollen, J.; Zuidema, W.; van Rooij, R.; Bockting, C. L. ChatGPT: Five Priorities for Research. *Nature* **2023**, *614* (7947), 224–226.
(3) Bender, E. M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*; FAccT '21; Association for Computing Machinery: New York, 2021; pp 610–623.
(4) Rillig, M. C.; Gould, K. A.; Maeder, M.; Kim, S. W.; Dueñas, J. F.; Pinek, L.; Lehmann, A.; Bielcik, M. Opportunities and Risks of the "Metaverse" For Biodiversity and the Environment. *Environ. Sci. Technol.* **2022**, *56* (8), 4721–4723.
(5) Strubell, E.; Ganesh, A.; McCallum, A. Energy and Policy Considerations for Deep Learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*; Association for Computational Linguistics: Florence, Italy, 2019; pp 3645–3650.
(6) Soga, M.; Gaston, K. J. Extinction of Experience: The Loss of Human–Nature Interactions. *Front. Ecol. Environ.* **2016**, *14* (2), 94–101.

https://doi.org/10.1021/acs.est.3c01106
Environ. Sci. Technol. 2023, 57, 3464–3466