

World view



By Kate Crawford

Generative AI is guzzling water and energy

First-of-its-kind US bill would address the environmental costs of the technology, but there's a long way to go.

Last month, OpenAI chief executive Sam Altman finally admitted what researchers have been saying for years – that the artificial intelligence (AI) industry is heading for an energy crisis. It's an unusual admission. At the World Economic Forum's annual meeting in Davos, Switzerland, Altman warned that the next wave of generative AI systems will consume vastly more power than expected, and that energy systems will struggle to cope. "There's no way to get there without a breakthrough," he said.

I'm glad he said it. I've seen consistent downplaying and denial about the AI industry's environmental costs since I started publishing about them in 2018. Altman's admission has got researchers, regulators and industry titans talking about the environmental impact of generative AI.

So what energy breakthrough is Altman banking on? Not the design and deployment of more sustainable AI systems – but nuclear fusion. He has skin in that game, too: in 2021, Altman started investing in fusion company Helion Energy in Everett, Washington.

Most experts agree that nuclear fusion won't contribute significantly to the crucial goal of decarbonizing by mid-century to combat the climate crisis. Helion's most optimistic estimate is that by 2029 it will produce enough energy to power 40,000 average US households; one assessment suggests that ChatGPT, the chatbot created by OpenAI in San Francisco, California, is already consuming the energy of 33,000 homes. It's estimated that a search driven by generative AI uses four to five times the energy of a conventional web search. Within years, large AI systems are likely to need as much energy as entire nations.

And it's not just energy. Generative AI systems need enormous amounts of fresh water to cool their processors and generate electricity. In West Des Moines, Iowa, a giant data-centre cluster serves OpenAI's most advanced model, GPT-4. A lawsuit by local residents revealed that in July 2022, the month before OpenAI finished training the model, the cluster used about 6% of the district's water. As Google and Microsoft prepared their Bard and Bing large language models, both had major spikes in water use – increases of 20% and 34%, respectively, in one year, according to the companies' environmental reports. One preprint¹ suggests that, globally, the demand for water for AI could be half that of the United Kingdom by 2027. In another², Facebook AI researchers called the environmental effects of the industry's pursuit of scale the "elephant in the room".

Rather than pipe-dream technologies, we need pragmatic

Within years, large AI systems are likely to need as much energy as entire nations."

Kate Crawford is a professor at the University of Southern California Annenberg, a senior principal researcher at Microsoft Research in New York City and author of the 2021 book *Atlas of AI*. e-mail: kate@katecrawford.net

The author declares competing interests; see go.nature.com/3woboth

actions to limit AI's ecological impacts now.

There's no reason this can't be done. The industry could prioritize using less energy, build more efficient models and rethink how it designs and uses data centres. As the BigScience project in France demonstrated with its BLOOM model³, it is possible to build a model of a similar size to OpenAI's GPT-3 with a much lower carbon footprint. But that's not what's happening in the industry at large.

It remains very hard to get accurate and complete data on environmental impacts. The full planetary costs of generative AI are closely guarded corporate secrets. Figures rely on lab-based studies by researchers such as Emma Strubell⁴ and Sasha Luccioni³; limited company reports; and data released by local governments. At present, there's little incentive for companies to change.

But at last, legislators are taking notice. On 1 February, US Democrats led by Senator Ed Markey of Massachusetts introduced the Artificial Intelligence Environmental Impacts Act of 2024. The bill directs the National Institute for Standards and Technology to collaborate with academia, industry and civil society to establish standards for assessing AI's environmental impact, and to create a voluntary reporting framework for AI developers and operators. Whether the legislation will pass remains uncertain.

Voluntary measures rarely produce a lasting culture of accountability and consistent adoption, because they rely on goodwill. Given the urgency, more needs to be done.

To truly address the environmental impacts of AI requires a multifaceted approach including the AI industry, researchers and legislators. In industry, sustainable practices should be imperative, and should include measuring and publicly reporting energy and water use; prioritizing the development of energy-efficient hardware, algorithms, and data centres; and using only renewable energy. Regular environmental audits by independent bodies would support transparency and adherence to standards.

Researchers could optimize neural network architectures for sustainability and collaborate with social and environmental scientists to guide technical designs towards greater ecological sustainability.

Finally, legislators should offer both carrots and sticks. At the outset, they could set benchmarks for energy and water use, incentivize the adoption of renewable energy and mandate comprehensive environmental reporting and impact assessments. The Artificial Intelligence Environmental Impacts Act is a start, but much more will be needed – and the clock is ticking.

1. Li, P., Yang, J., Islam, M. A. & Ren, S. Preprint at <https://arxiv.org/abs/2304.03271> (2023).
2. Wu, C.-J. et al. Preprint at <https://arxiv.org/abs/2111.00364> (2021).
3. Luccioni, A. S., Viguier, S. & Ligozat, A.-L. Preprint at <https://arxiv.org/abs/2211.02001> (2022).
4. Kaack, L. H. et al. *Nature Clim. Chang.* **12**, 518–527 (2022).