VAN BUSSEL
DOCUMENT SERVICES

The ecological footprint of Artificial Intelligence

Geert-Jan van Bussel

The emergence of artificial intelligence (AI) presents new opportunities for healthcare, communications, government and manufacturing. Despite growing concerns about AI-generated 'deep fakes', 'fake news' and other misinformation, the enthusiasm for this technology is staggering. But do we accept the impact of AI on our environment? Or its staggering demand for scarce resources such as energy and water?

The emergence of artificial intelligence (AI) has the potential to herald a new era of opportunity. The opportunities presented by AI are diverse and span a range of sectors, including healthcare, communications, government and manufacturing. The technology enables the creation of text, images, sound and art. AI can also help mitigate the effects of the climate crisis by developing intelligent energy grids, by developing infrastructures with little or no CO_2 emissions and by modelling climate predictions.

However, not all is positive. There are growing concerns that AI could play a role in the spread of misinformation, including ‘fake news’, ‘deep fakes’ and other forms of disinformation that could have a detrimental impact on our democratic society through populism and polarization.

A lesser-known consequence of AI is its impact on the natural environment. There is a substantial body of literature on this topic from recent years, but there is often a significant time lag between the publication of studies on this subject and their dissemination and uptake by the general public. The following examples illustrate this point.

In 2018, OpenAI reached the conclusion that the computing power required to train a large AI model had doubled every 3.5 months from 2012 onwards. The accuracy of results and time efficiency that can be achieved by harnessing the computing power of a vast number of computers in data centres necessitates a considerable amount of electricity. A significant proportion of data centres globally continue to rely, to some extent, on fossil fuels, resulting in a notable surge in CO_2 emissions.

In 2020, researchers at the University of Massachusetts conducted an analysis of several natural language processing (NLP) models and determined that the energy expenditure associated with training a single model resulted in CO_2 emissions of approximately 300,000 kg on average (equivalent to 125 return flights from New York to Beijing).

The training of ChatGPT-3 has been found to require the consumption of 1.3 gigawatt hours of electricity, resulting in the generation of 550,000 kg of CO_2. It is estimated by Bloomberg that the energy consumption necessary for training is only 40% of that required for operational purposes. Moreover, the training process necessitates the consumption of approximately 700,000 litres of water for the purpose of computer cooling. This quantity of water is equivalent to that which would be required by a nuclear power plant cooling tower.

In the year 2023, data centres operated by Google extracted a total of 24 billion litres of water from the environment. This represents a 14% increase compared to figures recorded in the previous year. In that

year, 20 billion litres of water were employed for the purpose of cooling. Two-thirds of this quantity was comprised of potable water. Furthermore, data from Microsoft's facilities indicate a 34% increase in cooling water consumption during the same period. In 2024, Microsoft's CO_2 emissions were 30% higher than in 2020, while Google's emissions increased by 48% over the past five years.

It is estimated that Sweden will witness a twofold increase in data centre electricity demand by 2030, while in Britain, AI is projected to necessitate 500% more energy over the subsequent decade. In the United States, data centres will account for 8% of all electricity consumption in 2030, representing a significant increase from the 3% they currently represent in 2022.

The advent of AI is an inevitable consequence of the progression of technology. However, there is a risk of scarcity in energy and water resources. This necessitates the identification of solutions that facilitate the advancement of environmentally sustainable initiatives. This necessitates transparency, as major technology companies have not been forthcoming about the outcomes of their efforts to reduce their ecological footprint. A number of measures, already implemented by major technology companies, can be intensified:

- The utilization of virtualization and clustering of virtual servers;
- The reduction of IT under-utilization: 30% of the world's servers are unused or under-utilized. This results in billions of dollars in costs and high CO_2 emissions from the electricity used. An under-utilized server still consumes 30-40% of the power it would at full capacity;
- The utilization of more energy-efficient chips using multi-core technology enables the maintenance of high performance whilst simultaneously reducing energy consumption. When applied in conjunction with a dynamic allocation of electricity based on real-time usage, this approach has the potential to yield significant gains;
- The utilization of elevated voltages in the distribution of electrical power enables the optimization of server efficiency;
- The implementation of enhanced cooling systems, based on other cooling fluids than potable water;
- Improved data centre locations, where natural resources can be employed in the generation of electricity or where it is feasible to generate and store substantial quantities of electricity.

Such measures (and others) are needed. Otherwise, initiatives to reduce emissions and water use will make little difference.

Most importantly, we need to be aware of when we, as users, are using generative AI and when we are not. Generative AI is a form of AI that creates content based on ‘prompts’ (user requests) to create and produce text, images, video and sound clips. This generative AI uses models with billions of criteria on which to generate. These are the most ecologically invasive.

Does using those models outweigh what you want to achieve with them? Do we have a goal? Or is it just for fun? Or because we can? These are questions that need to be asked beforehand if we want to use those models. A bit more awareness of the ecological effects of AI among users can’t hurt.

Whichever way you look at it, the ecological footprint has to go down. AI is already helping, but it can do much more with the scarce resources at its disposal. And scarce they are: do we want new homes or new businesses to have to wait years to be connected to the grid?

Perhaps we should unleash an AI prompt after all to solve this consumption problem....

Literature used:

1. D. Amodei, en D. Hernandez (2018). 'AI and compute', *Open AI Blog*, 16 mei. Online source, retrieved July 16, 2024: https://openai.com/index/ai-and-compute/#modern.
   Archive: https://archive.ph/B6Jdc#40%.

2. Dhar, P. (2020). 'The carbon impact of artificial intelligence', *Nature Machine Intelligence*, Vol. 2, Augustus, pp. 223-225. Online source, retrieved July 16, 2024: https://www.nature.com/articles/s42256-020-0219-9.

3. E. Farhat en W. Mathis (2024). 'AI will suck up 500% more power in UK in 10 Years, Grid CEO says', *Bloomberg Technology*, March 26, 2024. Online source, retrieved July 16, 2024: https://www.bloomberg.com/news/articles/2024-03-26/ai-will-suck-up-500-more-power-in-uk-in-10-years-grid-ceo-says.
   Archive: https://archive.ph/0a0mD#selection-1325.0-1325.64.

4. S. Ghaffary (2024). 'Big Tech's climate goals at risk from massive AI energy demands', *Bloomberg Newsletter*, July 11, 2024. Online source, retrieved Juli 16, 2024: https://www.bloomberg.com/news/newsletters/2024-07-11/big-tech-s-climate-goals-at-risk-from-massive-ai-energy-demands.
   Archive: https://archive.ph/BCKWu.

5. Goldman Sachs Research (2024). Generational Growth: AI, data centers and the coming US power demand surge, *Goldman Sachs*, April 28, 2024. Online source, retrieved July 16, 2024: https://www.goldmansachs.com/intelligence/pages/gs-research/generational-growth-ai-data-centers-and-the-coming-us-power-surge/report.pdf.
   Archive: https://web.archive.org/web/20240716075217/https:/www.goldmansachs.com/intelligence/pages/gs-research/generational-growth-ai-data-centers-and-the-coming-us-power-surge/report.pdf.

6. *Google Environmental report, 2023*. Online source, retrieved July 16, 2024: https://www.gstatic.com/gumdrop/sustainability/google-2023-environmental-report.pdf.
   Archive: https://web.archive.org/web/20240716233656/https:/www.gstatic.com/gumdrop/sustainability/google-2023-environmental-report.pdf.

7. *Google. Environmental report, 2024*. Online source, retrieved July 16, 2024: https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf.

8. H. Hodson (2023). 'Large, creative AI models will transform lives and labour markets', *The Economist*, April 22, 2023. Online source, retrieved July 16, 2024: https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work.
   Archive: https://archive.ph/tId0o.

9. U. Hölzle (2022). 'Our commitment to climate-conscious data center cooling', *Google Blog*, November 1, 2022. Online source, retrieved July 16, 2024: https://blog.google/outreach-initiatives/sustainability/our-commitment-to-climate-conscious-data-center-cooling/.
   Archive: https://archive.ph/086IM.

10. P. Li, J. Yang, M.A. Islam, en S. Ren (2023). 'Making AI less 'thirsty': Uncovering and addressing the secret water footprint of AI models'. *arXiv preprint*. Online source, retrieved July 16, 2024: https://arxiv.org/pdf/2304.03271.

11. Microsoft. *Environmental sustainability report*, 2022. Online source, retrieved July 16, 2024: https://news.microsoft.com/wp-content/uploads/prod/sites/42/2023/05/2022-Environmental-Sustainability-Report.pdf.
   Archive: https://web.archive.org/web/20240621093848/https:/news.microsoft.com/wp-content/uploads/prod/sites/42/2023/05/2022-Environmental-Sustainability-Report.pdf.

12. D. Patterson, J. Gonzalez, Q. Le, C. Liang, L-M. Munguia, D. Rothchild, D. So, M. Texier, en J. Dean (2021). 'Carbon emissions and large neural network training'. *arXiv pre-print*. Online source, retrieved July 16, 2024: https://arxiv.org/pdf/2104.10350.

13. R. Schwartz, J. Dodge, N.A. Smith, en O. Etzioni (2020). 'Green AI', *Communications of the ACM*, Vol. 63, No. 12, pp. 54-63.

14. E. Strubell, A. Ganesh, en A. McCallum (2020). 'Energy and policy considerations for modern deep learning research', C. Zheng, X, Fan, C. Wang, en J. Qi (red.). *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI*, Washington, No. 09, pp. 13693-13696. Online source, retrieved July 16, 2024: https://ojs.aaai.org/index.php/AAAI/article/view/7123.

15. A. Tapia Zafra en D. Gura (2024). 'AI wreaks havoc on global power systems', *Bloomberg Business*. Online source, retrieved July 16, 2024: https://www.bloomberg.com/news/articles/2024-06-24/microsoft-apple-amazon-ai-energy-needs-straining-ireland-us-malaysia-power-g. Archive: https://archive.ph/4QZEa