Energy Efficiency and Search

How AI is increasing the environmental costs of search engines

Provided by Digiconomist

https://digiconomist.net

Emerging technologies are increasing data center power demand

Worldwide data centers have been responsible for **~1% of global electricity consumption** between 2010 and 2018. Since 2018 we have witnessed the rapid expansion of electricity consumption for **cryptocurrency mining. Bitcoin** alone is currently responsible for ~**0.5% of global electricity consumption**.

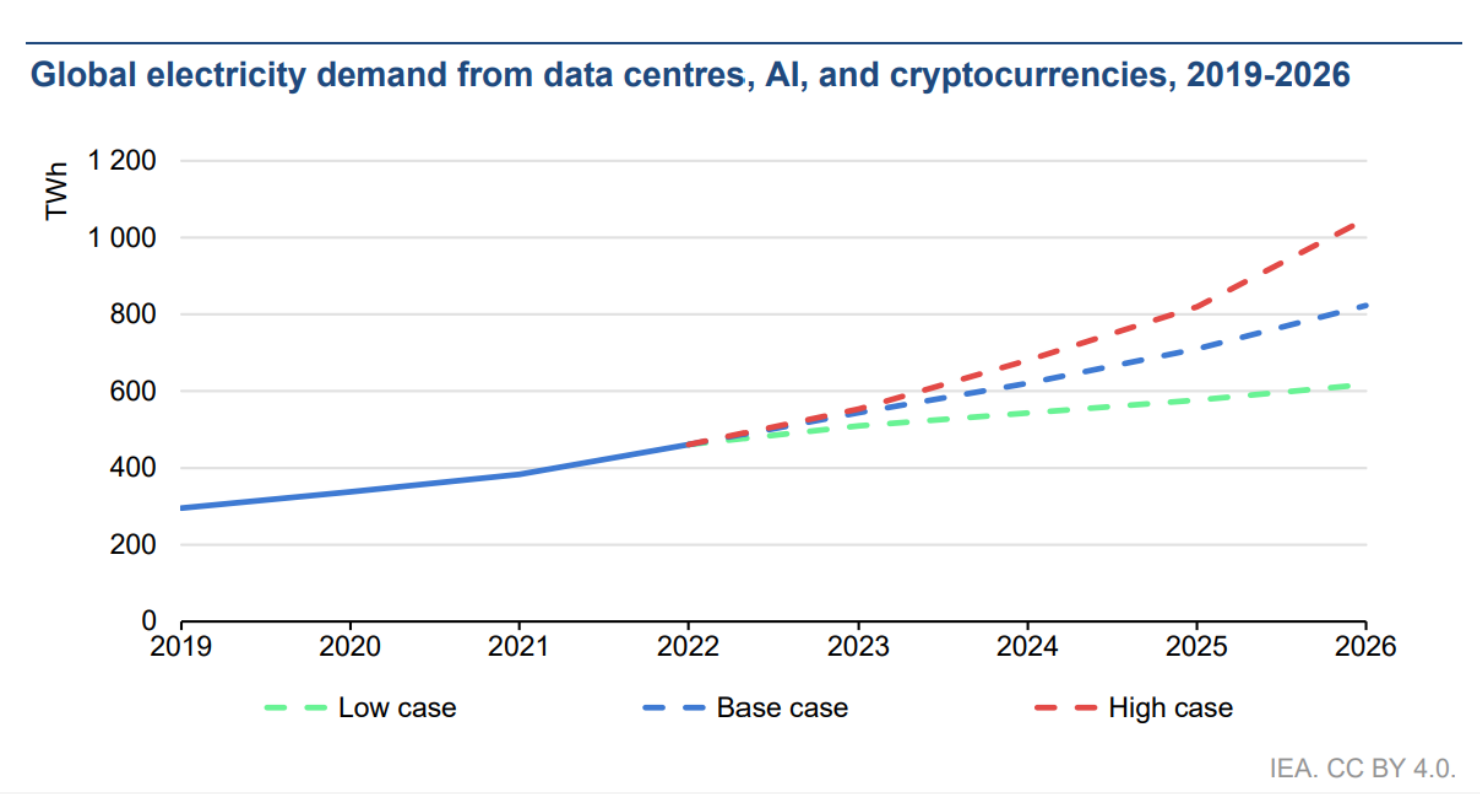
If not managed properly, **AI-related electricity consumption may experience a similar growth trajectory**.

In my own research *(Joule, 2023)* I found that AI could soon (by 2027) be **consuming more electricity than a country such as the Netherlands**. These estimates were already conservative at the time and the likely growth trajectory has been moving up.

It was recently revealed that OpenAI lobbied to open 5 to 7 massive 5GW data centers in the United States. **These data centers alone would consume as much electricity as Australia in total.**

Further increase in power consumption of data centers

The International Energy Agency (IEA) expects that in addition to AI, the power consumption of regular data centers, hardware and cryptocurrency mining will also increase further. This could make data centers responsible for up to **~3-4% of global electricity consumption.**



An example of how AI can make applications more energy-intensive

If we would turn Google search into ChatGPT today (assuming full adoption) using current available hardware and techniques, Google would consume as much electricity as a country such as Ireland just to power its search engine.

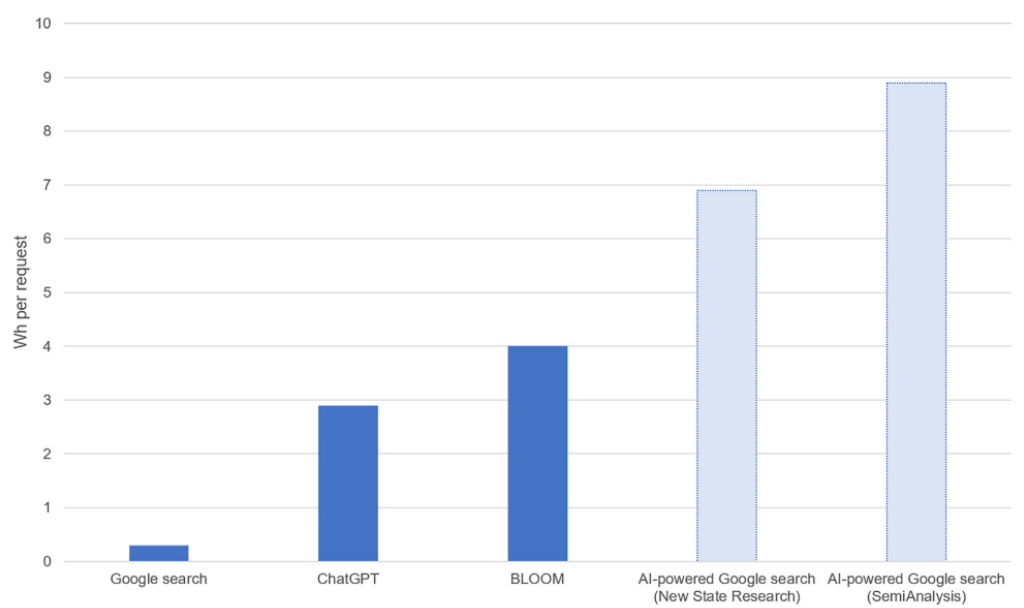
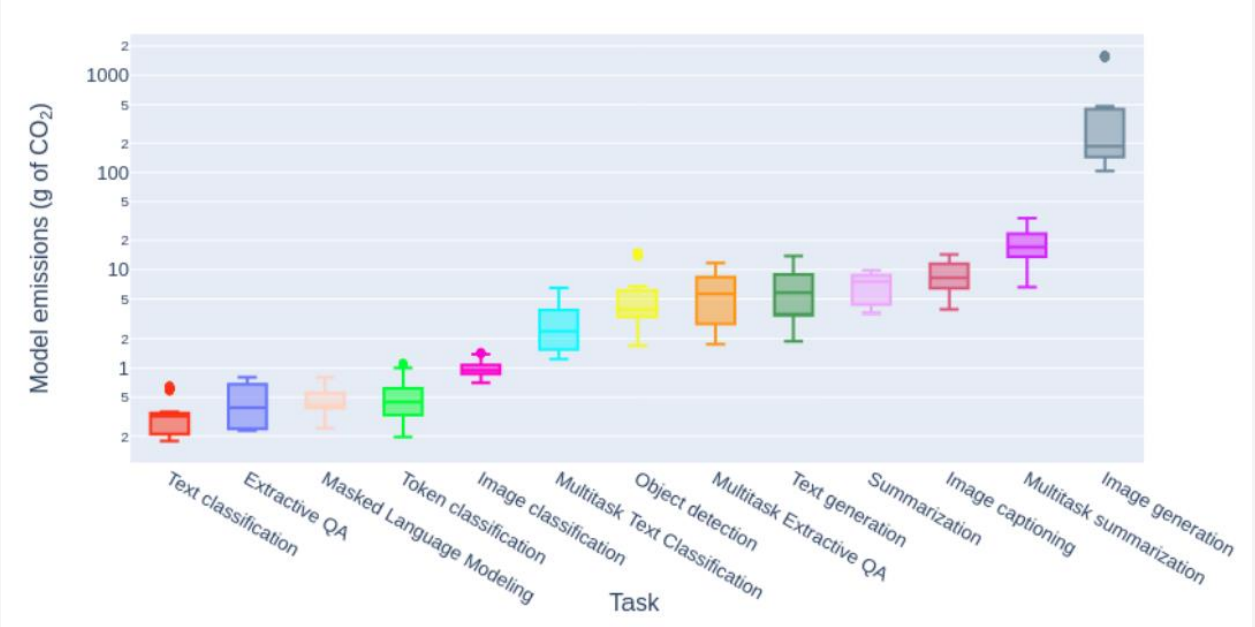
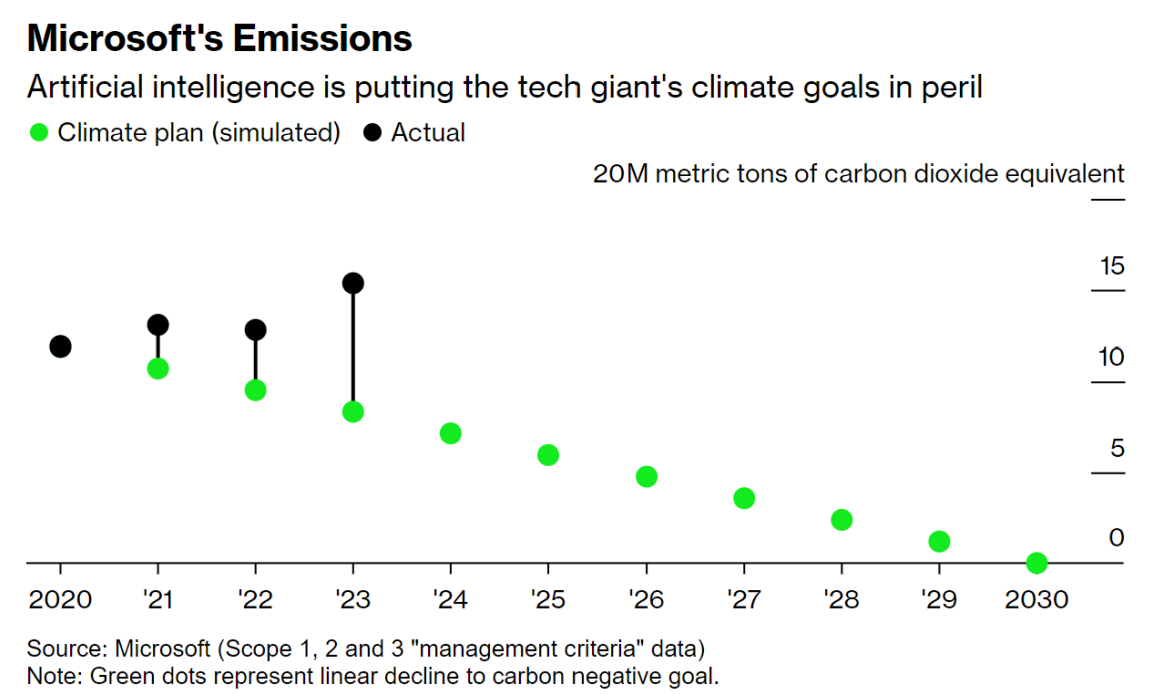


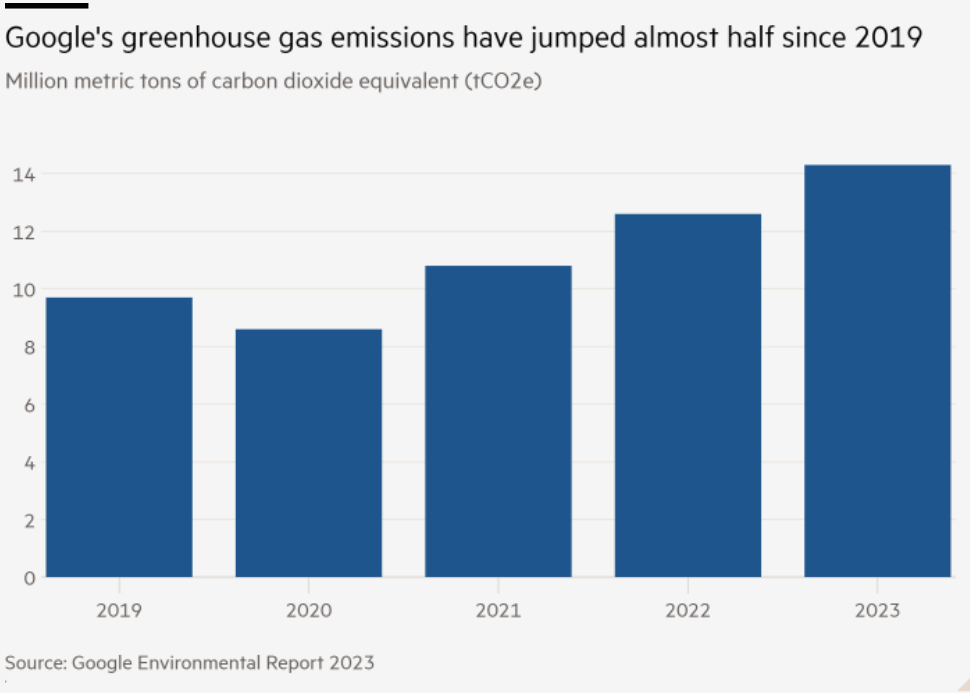
Figure 1. Estimated energy consumption per request for various AI-powered systems compared to a standard Google search



Current impact of the AI boom

"Microsoft's AI Push Imperils Climate Goal as Carbon Emissions Jump 30%"
"Google emissions jump nearly 50% over five years as AI use surges"





What makes AI energy-intensive technology

Artificial intelligence (AI) is an **umbrella term** that refers to a range of technologies and methods that enable machines to exhibit intelligent behavior. In the past year, we have seen several prominent examples of generative AI (another umbrella term) such as OpenAI's DALL-E and ChatGPT.

While these applications employ **distinct techniques**, they share a **common process**: an **initial training phase** followed by an **inference phase**. During training an AI model is fed **arbitrarily large datasets.** The model's initially arbitrary parameters are adjusted to align the predicted output closely with the target output. Following training, models are deployed into a production environment and begin the inference phase, where they generate outputs based on new data.

Bigger models are generally **more robust** and more robust models **perform better**. Simultaneously, bigger models require **more computational resources and electricity** to operate. It is this specific dynamic that makes AI energy-intensive.

We typically stop making models bigger as we reach a **tipping point** where the costs of makes a model bigger outweigh the (expected) benefit.

It is possible to make AI more efficient, but even these **more energy-efficient AI models seem destined to become bigger**, using up more data and resources. With more energy-efficient models we can afford to put in even more data and resources to further improve the models.

Local impacts

The previous slides covered the global impacts of data centres, but they also have significant impact on local communities, climate targets and electricity grids.

In Ireland, power consumption by data centers surpasses that of residential homes

AI boom: Enormous growth in US data centers

New US data centers are being built at the drop of a hat. It's not enough. If you need significant capacity, you have to rent years in advance.

Data centres curbed as pressure grows on electricity grids

Ireland, Germany, Singapore and China are imposing rules on building new server farms over energy usage fears

Amazon Tones Down Its Data Center Noise After Residents Sound the Alarm

After a year, the Great Oak neighborhood outside Manassas, Virginia, gets some relief.

Moving forward in a sustainable manner

We should keep in mind that AI is not a miracle cure. The experiences with ChatGPT clearly show that expectations regarding AI are too high. Ultimately, many end users will not benefit from the use of AI. A focus on these end users can already prevent a lot of waste. Solutions may not require the use of AI, or it may be possible to opt for a simpler rather than a more complex (and energy intensive) AI technique.

Furthermore, more awareness and transparency are now needed. Tech companies only provide less information about their AI applications, while detailed information is important for a good understanding of the power consumption of AI and to enable end users to deal more consciously with applications that use AI. The **AI Act** in Europe is a start, but it still leaves much to be desired. Environmental disclosure is on a voluntary basis, unlike e.g. environmental reporting for crypto assets in the new Markets in Crypto Assets Regulation (MiCAR). Other aspects of the AI Act may unintendedly affect the environmental impact of AI.

In any case, don't assume the benefits of AI will simply outweigh the costs. It is argued that AI could make certain goods and services more efficient, perhaps offsetting increased power consumption. However, the fact remains that price decreases for various goods and services can also lead to increased demand for them, ultimately increasing the overall demand for natural resources. This effect is known as the **Jevons paradox** and is evident throughout the history of technological change and automation.

Thank you

Contact info:

DIGICONOMIST

Digiconomist website (Digiconomist.net)

Digiconomist Twitter (@DigiEconomist)

Email: alex@digiconomist.net