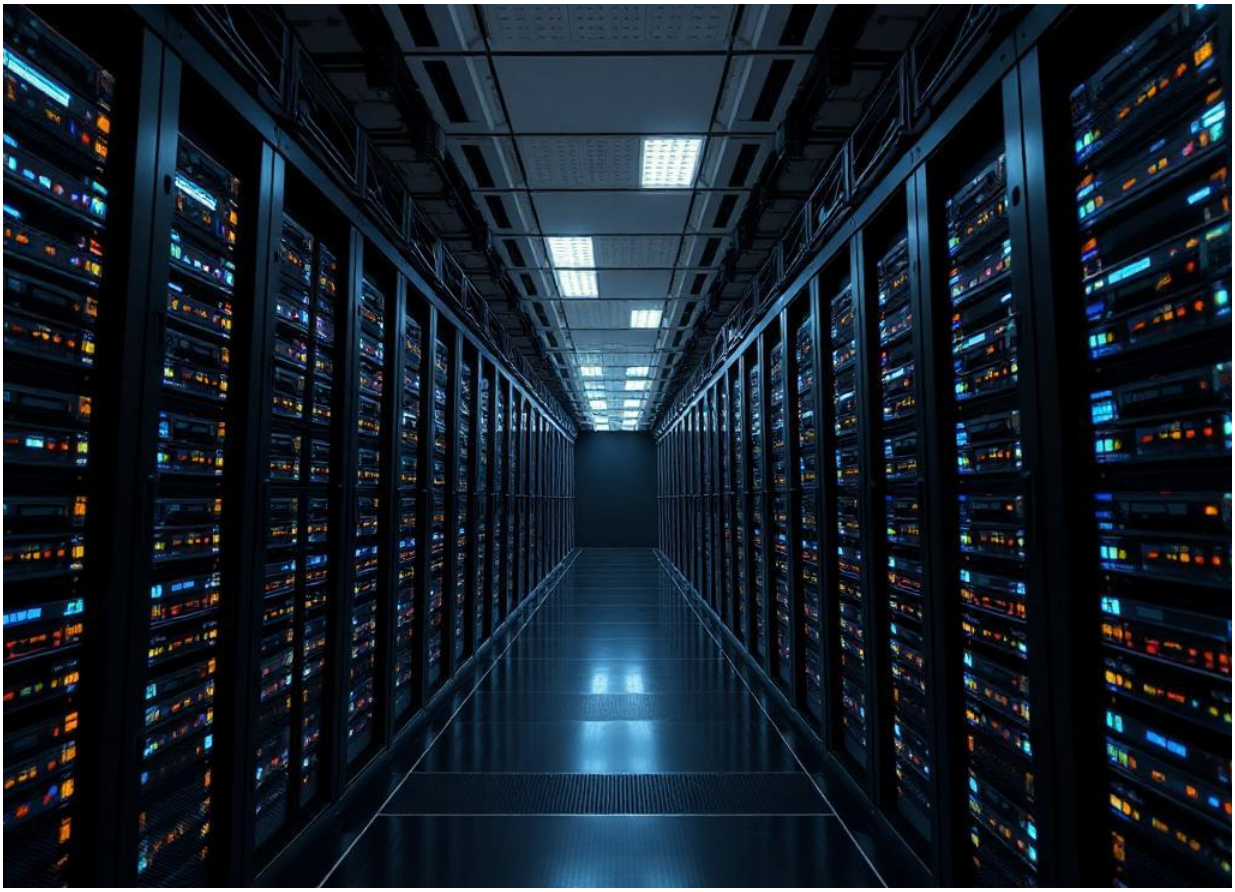

15 SEPTEMBER 2025

The Future Energy Challenge of AI in Data Centers

We are on the verge of a new revolution in the digital world: the exponential growth of AI and the associated pressure on global data centers. Whereas until recently we were concerned about the energy requirements of traditional servers, we are now facing a bigger challenge: the energy hunger of AI models. The amount of energy consumed by data centers worldwide is increasing at a rapid pace – driven by AI.

In 2024, data centers worldwide consumed about 415 TWh (terawatt hours), accounting for about 1.5% of the total electricity consumption on earth. AI now accounts for no less than 15% of this consumption, and this percentage will only increase. According to the International Energy Agency (IEA), global energy consumption by data centers could reach 945 TWh by 2030, representing nearly 3% of global electricity consumption.

But this is just the tip of the iceberg. AI servers, which are specifically designed for the heavy computational tasks of models like GPT3, consume as much as 30% more energy

annually than their traditional counterparts. While traditional servers show an annual growth of about 9% in energy demand, AI servers see an annual growth of 30%. This indicates the scale of the challenge ahead.

Training vs. Inference: the energy distribution

The development of large AI models such as GPT3 consumes an untold amount of energy. The training of such a model consumes more than 1287 MWh of electricity, with CO_2 emissions equivalent to driving 112 petrol cars annually. This is in stark contrast to the daily use of AI, or *inference*, for which the majority of energy is used: 60% of the total energy consumption goes to performing AI tasks that are done by millions of users worldwide.

For example, a single request to a model such as ChatGPT consumes an average of 2.9 Wh, compared to just 0.3 Wh for a Google search. This shows how AI significantly disrupts the energy balance in data centers and increases the challenge for the future.

Regional disparities and infrastructure pressures

The energy consumption of data centers is not evenly distributed across the world. In the US, for example, the share of data centers in total electricity consumption is expected to increase from 4% in 2023 to as much as 9 to 12% in 2030. A **160% increase** in data centre capacity is also expected in the UK, putting significant pressure on local infrastructure.

This also affects other critical resources, such as water, which is essential for cooling data centers. By 2027, global water consumption for data centers could reach as much as 6.6 billion m³ per year. This is already leading to restrictions on the expansion of data centers in some places, such as in Singapore and Dublin, where water shortages are hampering the construction of new facilities.

The comparison with dieselgate: lessons for the future

It may seem like an unusual comparison, but the energy challenge associated with AI can be compared to the fallout from the Dieselgate scandal. When Volkswagen was caught manipulating emissions tests, the long-term impact became clear: more than 53 million vehicles were sold worldwide with "defeat devices", which hid harmful emissions.

For example, the impact of AI on the world can also be invisible, but have enormous consequences. While Dieselgate mainly affected public health, the growth of AI data centers threatens to lead to structural problems in the areas of climate and energy security. Global energy consumption will not only increase, but also put pressure on infrastructure, which could lead to higher costs, disrupted ecological systems and even water shortages.

Sustainability initiatives and policymaking: the importance of transparency

In response to growing concerns about energy consumption in data centers, companies like Google are already starting to publish detailed emissions and energy data. For example, in 2023, it was reported that a single query for Google's AI model **Gemini emits** 0.03 g of CO_2, which is equivalent to the energy required for just 9 seconds of TV viewing. But there is a growing awareness that such figures may not tell the full story. Critical questions remain about the underreporting of indirect costs and the actual environmental impact.

Despite this, it is important that the industry remains committed to transparency and that standardized metrics are put in place to improve the energy efficiency of AI models. This should be supported by regional infrastructure planning that takes into account the burden that AI places on energy and water supplies.

Conclusion: what can we learn from dieselgate?

The AI revolution has undeniable benefits, but the environmental costs associated with the massive energy consumption of data centers cannot be ignored. As with Dieselgate, there is a high probability that we will suffer the consequences of this irresponsible energy consumption in the future.

It is crucial that we as a society proactively start regulating AI energy consumption, promoting energy efficiency, and supporting transparency initiatives. It would be irresponsible to allow this growth without intervention. We still have the opportunity to meet this challenge before we face the full impact of this invisible energy revolution.

Conclusion: The time to take action is now. Just as we saw with Dieselgate, the consequences of unregulated growth can be disastrous. AI must be developed responsibly, with full transparency and the necessary infrastructural adjustments. Only then can we find the balance between progress and sustainability.