Artificial Intelligence and Climate Change: Saviour or Contributor?

BEYOND URBAN LEGENDS: AI'S REAL ELECTRICITY DEMAND AND ITS CLIMATE OPPORTUNITY[1]

SEPTEMBER 2026

[1] The Vision team for this paper is Francesco Grillo and Linus Wendel

AI and climate change are two of the three great spectres haunting global policy, the third being the recurring threat of global warfare. They dominate the anxieties of governments, boardrooms, and voters alike, promising either to displace white-collar professionals or to render major European cities unliveable. Naturally, a fierce debate has emerged over how these two forces interact. Is artificial intelligence primarily an environmental villain, its insatiable hunger for electricity driving up emissions and roasting the planet? Or is it a saviour, offering hyper-precise forecasts to predict where climate disasters will strike next?

To cut through the ideological smoke hanging over the debate, which is a central focus of the upcoming global climate governance summit in Venice, several distinct trends emerge. Together, they suggest a surprisingly bullish outlook, provided society learns to harness the technology.

AI's Real Electricity Footprint:

The first hurdle in understanding AI is a problem of scale: human brains are notoriously poor at grasping exponential growth. Physical phenomena, including climate change, tend to move in single-digit percentages. The battle lines drawn at the 1995 Berlin climate summit aimed to cap global temperature rises at no more than 1.5°C to 2°C above pre-industrial baselines (around 13.8°C). The increase in global temperatures has, so far, followed a linear path. AI, by contrast, moves on a parabolic trajectory. Synthesising data from the International Energy Agency and S&P Global, the electricity consumed by training and inference has surged roughly tenfold[2] since November 30, 2022. The date OpenAI released ChatGPT and ignited the generative boom.

Yet user adoption and data processing have grown faster still. Alphabet's CEO, Sundar Pichai, recently noted that Google now processes 330 times more tokens per month than it did just two years ago. [3] The crucial economic

[2] International Energy Agency, "Energy and AI" (World Energy Outlook Special Report), 2025, cross-checked against S&P Global Commodity Insights electricity-demand tracking.
[3] Sundar Pichai, Google I/O 2026 keynote, 20 May 2026: monthly tokens processed across Google's products rose from 9.7 trillion in April 2024 to over 3.2 quadrillion in May 2026, an approximately 330-fold increase.

takeaway is that while total power draw has exploded, the energy cost per token has plummeted.

So perspective is required regarding AI's actual carbon footprint. While a tenfold increase in power consumption sounds alarming, it brings total AI electricity demand to roughly 150 terawatt-hours (TWh) per year. The total electricity consumption of servers makes up roughly 1.5% of global consumption but only between 30-45% of server capacities are used for AI. This means that AI uses only about 0.5% of global energy consumption which is roughly equivalent to the annual power consumption of Egypt. [4] When considering global energy consumption, which also accounts for oil, coal and non-grid power the share of AI falls to just 0.1%. For comparison, the single existing heavy industry aluminium smelting, already consumes roughly 3.5% of global electricity on its own[5], more than double the entire data centre sector combined.

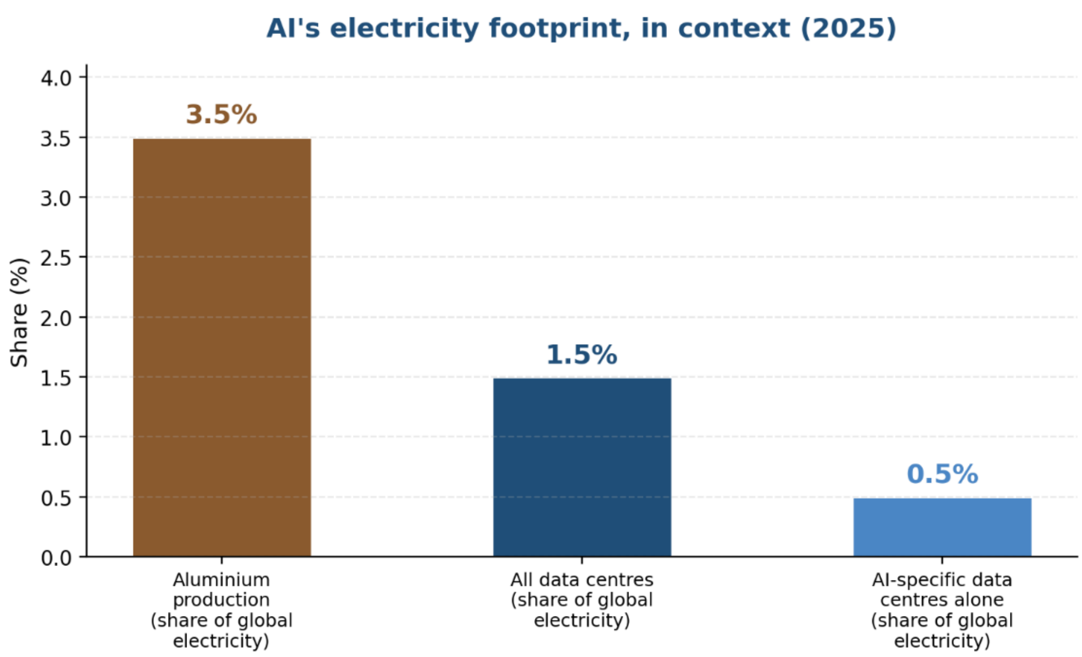

Figure 1. AI's electricity footprint in comparison to aluminium production and all data centres (Vision based on data from IEA, OWID and Calzadilla et al.)

Yet the aggregate figure obscures a more localized problem. What strains a grid is not how much electricity AI consumes worldwide, but how much it demands

[4] Our World in Data (Hannah Ritchie), "How much electricity does AI consume?" 2025–2026

[5] Calzadilla, A., et al. (2026). "The dynamic global aluminium use across production systems to consumption systems," Ecological Economics, Vol. 239

in one place, all at once. In Memphis, Tennessee, xAI's Colossus data centre could not wait for a conventional grid connection and instead built what amounts to its own off-grid power plant: as many as 35 methane gas turbines, several installed without air permits, together generating some 400–500 megawatts. Compare that with the data centres now clustering across Iceland and the Nordics, where Microsoft, Google and others have committed billions of dollars precisely because the local grid can absorb them without strain: Iceland's electricity is 100% renewable, split between hydro and geothermal,[6] and Norway draws roughly 98% of its power from hydro alone.[7] The lesson is not that AI's electricity demand is secretly enormous; it is that where that demand lands matters as much as how large it is.

AI's Opportunities for Climate Adaption:

Meanwhile, processing efficiency is yielding massive dividends in predictive power, particularly for chaotic climate systems. The European Centre for Medium-Range Weather Forecasts (ECMWF) has rolled out its AI Integrated Forecasting System (AIFS), which uses up to a thousand times less energy and is up to 10 times faster than traditional physics-based models.[8] It cannot pinpoint a hailstorm on a specific street corner months in advance, but it dramatically extends lead times and tightens spatial resolution. Paired with quantum computing, AI could eventually allow policymakers to model the exact economic costs of climate change for specific regions, alongside the most cost-effective mitigation strategies.

The Importance of Optimizing AI Use:

How AI is implemented into workflows, research, and private inquiries is another important issue. Currently models that require large quantities of tokens are tasked with solving problems that require much less computing power. For instance, using Anthropic's Fable 5 for a simple translations is both unnecessary and wasteful. Artificial Analysis found that "thinking" models can

---

[6] National Energy Authority of Iceland

[7] Norwegian Water Resources and Energy Directorate (NVE)

[8] European Centre for Medium-Range Weather Forecasts, press release, 1 July 2025: "First openly available ensemble model using Artificial Intelligence goes live – delivering forecasts over 10 times faster and cutting energy usage by 1000."

use up to 20x more tokens[9] compared to "non-thinking" ones. Using the right tool for the right task is therefore crucial.

However, due to the novelty of AI and the lack of proper price signals (many AI subscriptions do not charge by the token, and companies encourage employees to use as many tokens as possible regardless of the task to speed up AI integration) AI is used in a wasteful manner. It is the equivalent of using 1200L of water for a shower when 60L (the average) is more than enough.

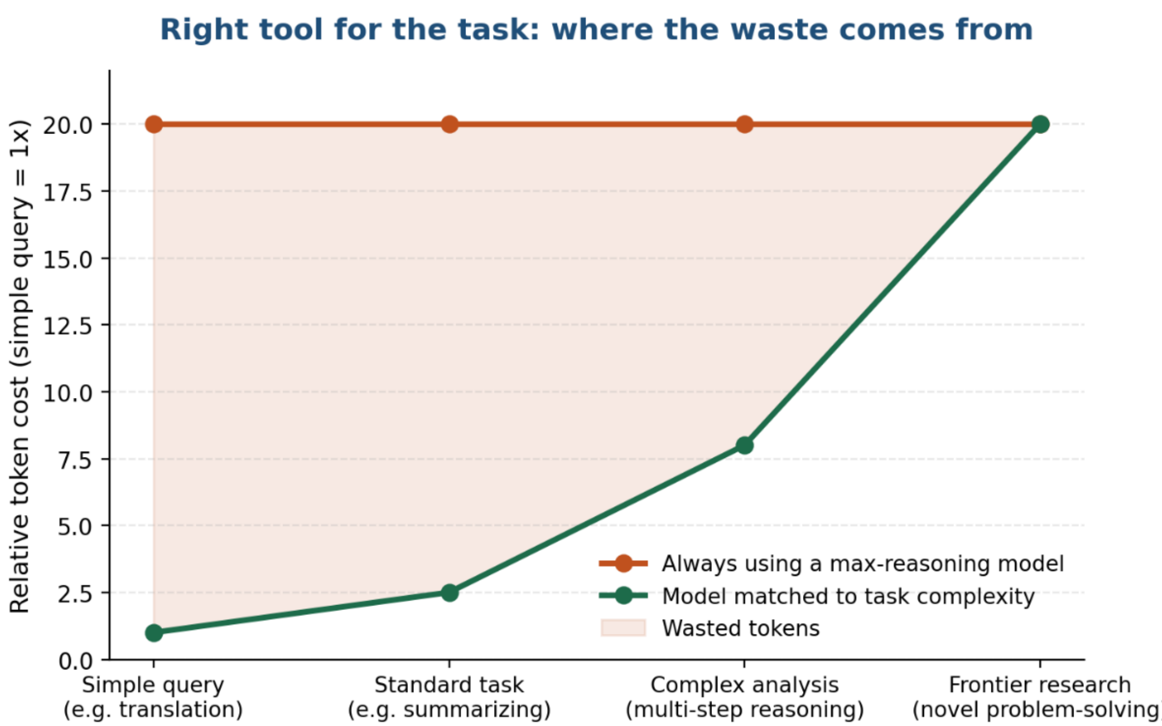

Figure 2. Defaulting to a max-reasoning model regardless of task complexity wastes tokens (Vision on Artificial Analysis data)

Diminishing Utility of More Data and Longer "Thinking":

None of this eliminates the downside risk. Pretraining itself is approaching a wall: frontier labs will soon exhaust the reservoir of high-quality human-generated text. The industry's answer has been to relocate the scaling problem from training to inference. This means letting models "think" longer each time a prompt is submitted rather than making them permanently larger. This is precisely why reasoning-enabled models consume so many more tokens than their non-reasoning predecessors: the compute that used to be spent once,

[9] Artificial Analysis, "How many tokens do reasoning models use vs. non-reasoning?" 16 April 2025.

during training, is now spent repeatedly, on every query. The danger is that aggregate demand for compute keeps climbing even as the returns on thinking harder (like the returns on training bigger before) begin to flatten.

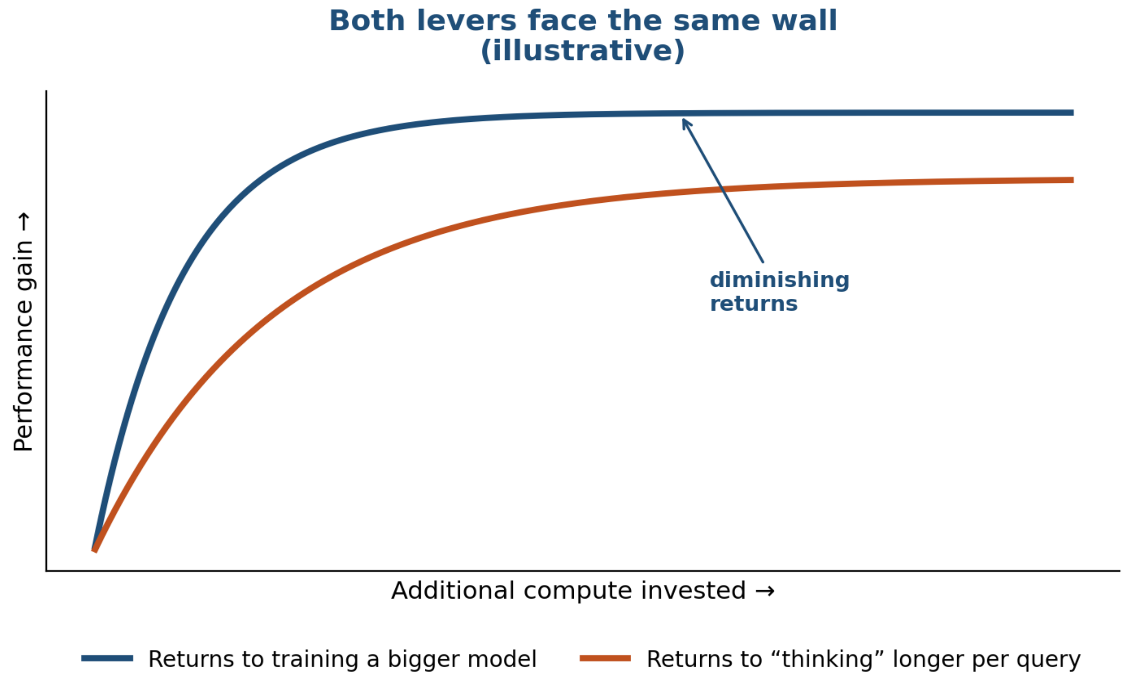

Figure 3. Both industry's scaling levers (bigger training runs and longer inference-time "thinking") produce steadily smaller gains per unit of compute (illustrative).

Summary:

Taken together, these trends suggest neither blind optimism nor reflexive alarm is warranted. AI's footprint is real but proportionate; its forecasting dividends are already measurable; and a large inefficiency is wasted tokens, not wasted electricity, which is a solvable problem of incentives rather than physics. The technology will not save the planet on its own, nor will it roast it. Whether it tips one way or the other depends less on the models themselves than on the institutions meant to govern them. To manage the strain on grids, enforce proportionate prices for tokens and manage run-away demand for a product that faces walls in terms of marginal utility gains. Yet unlike climate change, which has its annual UN summits, no international institution currently exists to police the machines. New forms of global governance are therefore required to tackle this issue.