

A brief on the environmental footprint of artificial intelligence

The Cloud Has a Body

WHAT
AI'S
CLIMATE
STORY
LEAVES
OUT



SEPTEMBER 2026



This report was prepared by Effective Thoughts, in collaboration with SMCSE Solutions and Wisdom Inn for Skill Enhancement (WISE), to contribute to the global discourse on artificial intelligence and its environmental consequences.



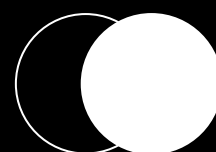
Effective Thoughts is a knowledge platform dedicated to producing and disseminating research-driven, reason-based, and human-oriented content across economics, technology, and society.



SMCSE Solutions is an IT, marketing, and business consulting firm offering digital transformation, MarTech, and data & AI intelligence to organizations worldwide.



Wisdom Inn for Skill Enhancement is a leadership and managerial coaching company specializing in tailored life and business programs for social, educational, and economic development.





This report is published under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). You may share it in full, with attribution, for non-commercial purposes. Commercial use, adaptation, or republication of excerpts requires written permission.

Third-party material reproduced or adapted in this report remains the property of its rights holders.

Suggested citation: Ashhar, M. and Raja, H. B. (2026). *The cloud has a body: What AI's climate story leaves out*. Effective Thoughts.

Ashhar conceived and directed the study and led revisions; H. B. Raja conducted the research and drafted the report.

LEAD AUTHORS

Muhammad **Ashhar**
Editor-in-Chief



Hijab Bano Raja
Economic Analyst





People talk about how much energy it takes to train an AI model – but it also takes a lot of energy to train a human. It takes about 20 years of life – and all the food you consume during that time – before you become smart.

— SAM ALTMAN, CEO OF OPENAI [1]

TABLE OF CONTENTS

1. What AI's Climate Story Leaves Out	6
2. What Global AI Infrastructure Looks Like	7
3. What the Cloud Costs	10
3.1 Energy	10
3.2 Water	11
3.3 Minerals and e-waste	11
4. The Rules Are Still Catching Up	12
4.1 The European Union: transparency without limits	12
4.2 The United States: fragmented rather than comprehensive	14
4.3 A side-by-side comparison	16
5. Where the Tradeoffs Get Real: The Singapore–Johor–Batam Corridor	17
6. What Doing It Better Looks Like	21
6.1 Companies	21
6.2 Countries	24
6.3 Efficiency is not the same as reduction	25
7. What Responsible Adoption Requires	26
7.1 Consumers and enterprise buyers: Driving demand for transparency	26
7.2 Industry: Advancing efficiency while strengthening accountability	27
7.3 Governments: Establishing clear and enforceable regulatory frameworks	27
7.4 International cooperation: Closing the global governance gap	28
7.5 Looking ahead	29
8. References	30





What AI's Climate Story Leaves Out

Artificial intelligence is increasingly marketed as a tool for environmental good. It tracks methane leaks from space, monitors biodiversity loss across remote ecosystems, and sharpens the climate models that governments use to plan for a warmer world [2]. That is true. But this story is also incomplete.

When Sam Altman told an audience that “people use a lot of water,” he was making a defensible point: AI does not, in aggregate, use more water or electricity than agriculture or heavy manufacturing today, globally. But critics were quick to note that this framing dodges the actual question.

The issue, one can argue, is not whether AI consumes environmental resources, but whether the explosive, concentrated growth of AI infrastructure globally can exert pressures on local water and energy systems, perhaps more so in some regions than others, faster than current regulatory frameworks can track or manage.





What Global AI Infrastructure Looks Like

Before totaling up the cost, it helps to see the scale and shape of the thing being costed. “The cloud” is not a diffuse, placeless entity. It is a specific, unevenly distributed set of physical buildings.

The infrastructure that powers AI (the servers, processors, and storage systems) lives in data centers. There are more than **11,000** data centers operating worldwide, across roughly 174 countries [4]. That number is growing fast: industry forecasts expect it to pass **12,000** by the early 2030s, alongside close to **\$7 trillion** in cumulative global investment in new and upgraded facilities by 2030 [5]. But “worldwide” is doing a lot of work in that sentence; the footprint is heavily concentrated in a small number of places.

The United States alone hosts more than **4,400** data centers, more than the next several countries combined [3]. Germany and the United Kingdom follow, each with just over **500**. China’s site count looks comparatively modest by this measure (around **370**), though that undercounts its actual computing capacity.¹ Chinese facilities tend to be built at a larger average scale, and government estimates of total operating sites run considerably higher than commercial trackers report [6].

¹ Note: counts vary by provider and definition

A second-tier market (Canada, France, the Netherlands, Australia, Japan) each host a few hundred, and a third tier of fast-growing markets (India, Singapore, Malaysia, Brazil, Ireland) is smaller in raw count but growing at some of the highest rates anywhere.

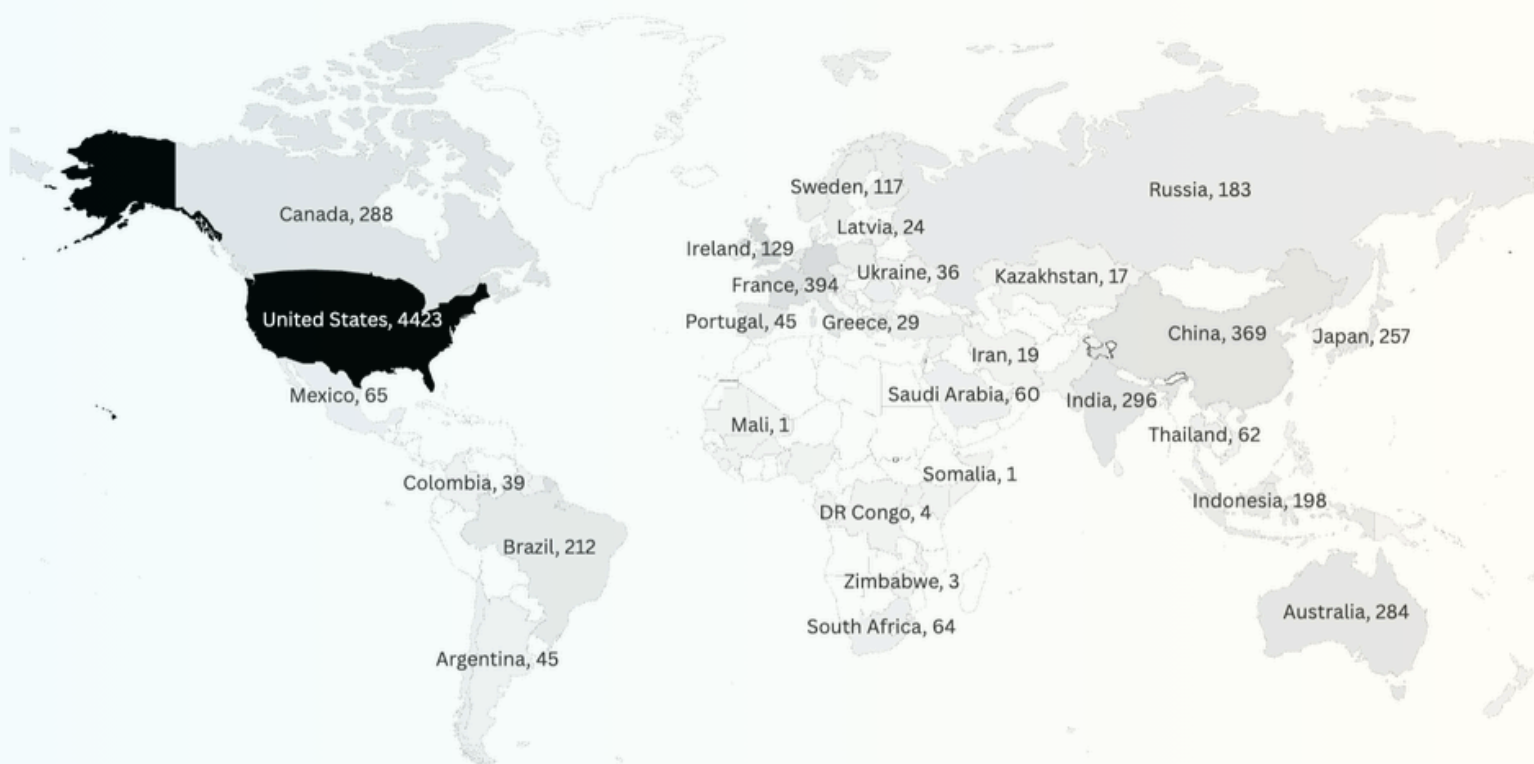


FIGURE 1. Global concentration of operational data centers by country, 2026. Some data points are hidden for visual effects. Data source: [Statista](#) [3]

Three things about this map matter for what follows. First, a handful of regions (Northern Virginia, Ireland, the Nordics, Singapore–Johor) absorb a disproportionate share of the water and electricity strain, which is why national, aggregated statistics can understate the pressure felt in any one of them.

Second, growth is now fastest precisely where infrastructure is thinnest and grids are least prepared; Southeast Asia, South Asia, Latin America, and the Gulf are all adding capacity faster than legacy markets, often without the regulatory or grid maturity those legacy markets eventually built up over decades.

Third, this map depicts general-purpose data centers, not AI-specific capacity. But AI workloads are increasingly what's driving new construction everywhere on it, and the newest AI-optimized facilities draw far more power per rack than the internet infrastructure built over the previous two decades [7]. The next section turns to what this infrastructure costs to run.





What the Cloud Costs

There is a physical infrastructure behind every AI query: server racks, cooling systems, transmission lines, and the mines that supply the minerals inside every chip. AI's environmental footprint is most clearly understood through three interconnected dimensions: its demand for energy, its consumption of water, and its reliance on resource-intensive hardware and critical minerals.

3.1 Energy

Global data center electricity demand is on track to **more than double** by 2030, reaching roughly **945** terawatt-hours, more than Japan's entire current electricity consumption [7]. In the United States, the effect is starker still: by 2030, American data centers are expected to consume more electricity than the country's aluminum, steel, cement, and chemical industries combined [7]. For a sense of scale at the individual level, a single ChatGPT query is estimated to use around **ten times** the electricity of a standard Google search [8].

This is not the whole picture, though. Research shows that efficiency per computational task is improving faster than at almost any point in energy history. It also finds that renewables are the fastest-growing power source for data centers: they supplied about **27%** of data center electricity in 2024 and are projected to grow at an average rate of roughly **22%** a year through 2030 [9]. Efficiency gains are real. Whether they are



sufficient to offset AI's growth remains an open question this brief returns to later.

3.2 Water

Global AI-related water withdrawals are projected to reach **4.2** to **6.6 billion** cubic meters a year by 2027, which is **four** to **six times** Denmark's total annual water withdrawal [10]. In Texas alone, data center water use is estimated at roughly **25 billion** gallons annually as of early 2026, with projections rising to between **29** and **161 billion** gallons by 2030 [11].

Cooling towers are often the most visible part of a data center's water use. But roughly **60%** of a typical data center's water footprint is indirect: it is the water consumed to generate the electricity the facility draws from the grid, not the water used on-site for cooling [12].

Many of the most-cited water figures (including the "six times Denmark" comparison) trace back to a 2023 academic estimate that has since been repeated across policy reports, journalism, and industry commentary [10]. The figures still hold up. What's likely immature is how this kind of data gets measured and tracked.

3.3 Minerals and e-waste

The environmental cost of AI begins before a single server is switched on. Producing a single two-kilogram computer requires roughly **800** kilograms of raw material extraction [13]. Chip manufacturing depends on a concentrated set of critical minerals (gallium, germanium, copper, nickel, cobalt, and rare earths) where supply is geographically concentrated and geopolitically sensitive.

China holds a near-monopoly on gallium and roughly **70%** of global germanium supply [14]. Indonesia, meanwhile, already dominates global nickel production, a mineral central to batteries and electronics. AI is not a passive bystander to this supply chain; it is actively straining it. AI-specific demand alone is projected to push gallium consumption up **85%** and germanium consumption up **37%** by 2033, layering fresh pressure onto mineral markets that were already concentrated and geopolitically fragile before AI accelerated demand [15].





The Rules Are Still Catching Up

We begin with the European Union's regulatory picture, then turn to the United States.

4.1 The European Union: transparency without limits

The EU's environmental governance emerges from a few different overlapping instruments.

The AI Act (2024) is, at its core, a risk and safety law, not an environmental one. But it does something relevant: providers of general-purpose AI models must document the computational resources and energy consumed during training and, where known, during testing and fine-tuning [16]. This is the first legal requirement anywhere for AI developers to quantify part of their environmental footprint.

The gaps, however, are noted across the literature. Unlike one-time costs of training a model, the Act does not regulate inference energy (electricity consumed every time a user interacts with it once deployed). It does not require lifecycle emissions reporting, carbon-intensity targets or renewable-energy obligations, and water-use standards.

Researchers studying this gap argue that inference may eventually exceed training as the dominant source of AI's energy footprint, which makes its absence from the law a weakness [17].



The Energy Efficiency Directive (EED, Directive (EU) 2023/1791) is arguably the EU's strongest instrument affecting AI infrastructure, even though it was not written with AI specifically in mind. It requires operators of data centers above **500kW** of installed IT power to report standardized sustainability indicators, such as electricity consumption, Power Usage Effectiveness (PUE), Water Usage Effectiveness (WUE), renewable energy share, and waste-heat recovery, to a European database [18].

The goal, for now, is transparency and letting operators be compared against each other, not mandatory performance standards. The European Commission is developing an EU-wide data center rating scheme and has signaled that transparency alone may prove insufficient as AI-driven electricity demand accelerates [19].

The Corporate Sustainability Reporting Directive (CSRD) separately requires large companies to disclose Scope 1 (direct), Scope 2 (from purchased electricity and energy), and Scope 3 (indirect value-chain) emissions. Scope 3 covers emissions the company doesn't own or directly control and can make up **70–90%** of the organization's total carbon footprint [20]. Even though the CSRD was not designed as an AI regulation, for hyperscale cloud providers, this indirectly captures an important share of AI-related environmental impact.

These are meaningful steps. Still, reporting does not automatically control consumption. A facility can report high electricity and water use while remaining fully compliant. Moreover, current rules are weighted toward the training phase of AI development rather than the deployment phase, even though inference could likely become the larger long-term cost.

Most tellingly, even mandatory disclosure has been quietly watered down. More data centers are now required to report under the EED, but the actual numbers for each facility are increasingly kept confidential. Only **36%** of EU data centers (**770** out of **2,161**) reported their data [21].

Microsoft, Google, Amazon, Meta and Netflix, as part of industry groups, reportedly lobbied “to hide the environmental toll of their data centers” [22].

4.2 The United States: fragmented rather than comprehensive

The US has no federal equivalent to the AI Act for environmental sustainability. Governance instead is spread across executive policy, voluntary technical standards, corporate disclosure regimes, and a growing patchwork of state law.

In 2023, a **federal executive order on AI safety** told agencies to study how AI infrastructure affects energy systems and critical infrastructure, but it never required AI developers to report their environmental impact. That entire order was scrapped within hours of the new president taking office in January 2025 [23], and a new executive order that July went the opposite direction: it told agencies to speed up environmental approvals for AI data centers, including exemptions from standard environmental review and easier access to federal land [24].

While the EU has been adding disclosure requirements, the US's biggest recent move has been to remove review steps.

The National Institute of Standards and Technology's **AI Risk Management Framework** has become the dominant US reference point for responsible AI, but environmental sustainability is not one of the framework's seven core 'trustworthy AI' characteristics (valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, and fair). It appears only as a single subcategory (**MS-2.12, 'environmental impact and sustainability'**) buried inside the framework's Measure function [25].

Separately, the **Securities and Exchange Commission (SEC)** adopted climate-disclosure rules in 2024 that would have required public companies to report climate risk where it is financially material (potentially capturing data center electricity demand and emissions), but the rules were stayed before ever taking effect and the SEC is now moving to rescind them entirely [26]. Some large companies still report this information voluntarily or to satisfy state-level rules, but such disclosures are company-wide, not AI-specific, and vary widely in depth.

While the EU has been adding disclosure requirements, the US's biggest recent move has been to remove review steps.

The real US regulatory activity is happening at the state level, and it is happening fast. More than forty states considered **267** data center-related bills in 2025 alone [27], and over **300** additional bills were introduced across 30 states in just the first six weeks of 2026 [28]. Some of this legislation has real teeth:

- Minnesota created a dedicated water-permitting requirement that allows denial where a project threatens watershed health;
- Utah now requires disclosure of a facility's water source;
- Oregon's POWER Act requires large data center operators, not residential ratepayers, to cover the cost of new grid infrastructure they require [29].

The result is a country moving in two directions simultaneously: **a federal government actively removing environmental review and a growing number of states moving to add it back.** The EU has built the first serious legal architecture for AI-environmental transparency and is already watching parts of it get quietly hollowed out by industry lobbying, while the US has almost no federal architecture at all and is instead watching an active but uncoordinated experiment play out state by state.

4.3 A side-by-side comparison

We summarize the EU and the US regulatory picture in Table 1.

TABLE 1.
Regulatory approach to AI's environmental footprint, EU vs. United States (2025–26)

ISSUE	EUROPEAN UNION	UNITED STATES
AI-specific law	Yes — AI Act (2024)	No federal equivalent
Data center reporting requirement	Mandatory for centers over 500kW	No federal mandate – state rules only
Energy reporting	Yes, mandatory	No federal requirement – a few states only
Water reporting	Yes, mandatory	Limited, state-by-state
Renewable-energy requirements	Emerging — rolled out country by country	No federal requirement
Carbon reporting	Required, covering all three emissions scopes	General corporate disclosure only
Facility-level data actually public	Collected, but kept confidential	No federal requirement
2025–26 direction of travel	Tightening (rating scheme in development)	Federal: loosening. States: tightening





Where the Tradeoffs Get Real: The Singapore–Johor–Batam Corridor

AI is often discussed as weightless: purely virtual technology. But that framing obscures the fact that AI runs on physical hardware tied to electrical grids and water supply. Its environmental impacts are intensely local, and one of the clearest places to see this is the Singapore–Johor–Batam corridor in Southeast Asia, where AI-driven demand is reshaping cross-border resource politics in real time.

Singapore remains the region’s digital command center, but it faces constraints on land, electricity, and water. Between 2019 and 2022, the government paused approval of new data centers altogether over exactly these concerns. When approvals resumed, they came with conditions attached: new facilities had to demonstrate best-in-class power usage effectiveness (PUE), secure renewable energy, and meet explicit efficiency targets under Singapore’s Green Data Centre framework [30].

The three-year pause in Singapore, combined with constraints and conditions, consequently prompted investors to look beyond the shore. AI infrastructure growth did not stop; it moved. Much of the region’s new hyperscale capacity has

instead been built just across the border in Johor, Malaysia, whose operational data center capacity roughly doubled from about **401 MW** in 2023 to around **897 MW** in 2024, with several additional gigawatts in the development pipeline [31]. Hyperscalers favor Johor for straightforward reasons: cheaper land, abundant industrial zoning, lower electricity costs, and proximity to Singapore's financial and technology ecosystem [32].

Per environmental economics, this pattern arguably shows **carbon leakage**, where regulatory pressure in one jurisdiction relocates environmental cost to neighboring countries rather than eliminating it. Singapore increasingly captures the high-value layer of the AI economy (headquarters, finance, cloud services, and AI startups), while Johor hosts the physical layer: warehouses, substations, cooling infrastructure, and the electricity and water demand that comes with it. Whether Singapore has become "AI-light" and Johor "AI-heavy" is now a live question in regional policy circles, and it reframes the entire debate: the benefits of AI's growth are increasingly regional, but so, unevenly, are its costs.

Water sharpens the point further. Johor has historically had greater water availability than Singapore, but southern Malaysia has faced droughts and reservoir shortages – as recently as April 2026, when a prolonged heatwave pushed three major Johor dams below critical levels, affecting roughly a million residents [33] – and rising industrial demand of its own, even as Singapore continues to import treated water from Johor under the 1962 Water Agreement, a bilateral arrangement now over six decades old [34].

Malaysia's own water regulator has found that of over **800 million** liters a day requested by proposed data center projects across three states, only about **142 million** liters can realistically be supplied, with Johor alone accounting for roughly **90%** of that shortfall [35]. This is not yet a full-blown crisis, but it is an emerging governance problem inside an already tightly interconnected regional water system, precisely the kind of issue that current national frameworks are not built to manage, since the infrastructure is regional, but the water treaty and the permitting regime are not.

Batam, Indonesia (the corridor's third node) is also emerging in the regional AI ecosystem. Sitting only 20 kilometers from Singapore's south coast and carrying more than **80%** of

Indonesia's international internet traffic [36], the island has attracted major hyperscale investment since the early 2020s, with the pace accelerating sharply since 2025. This includes a **\$5 billion** AI data center project from the China-backed Equator Gate System [37], alongside earlier builds from DayOne (underway since 2021) and NeutraDC [38]. Like Johor, though, Batam's fast growth is testing the same energy and water limits playing out across the rest of the corridor.

There is a renewable-energy mismatch running underneath all of this, too. Hyperscalers frequently advertise "100% renewable energy" commitments, but the underlying regional grids remain fossil-heavy: coal and gas together made up roughly **79%** of Malaysia's electricity generation in 2023 [39], and Indonesia's reliance on coal has actually increased, from **56%** of generation in 2015 to nearly **67%** in 2024 [40].

A renewable-energy purchase agreement does not change what physically comes out of the wall at a given substation; it is a financial and accounting instrument layered on top of a grid that, in this region, is still substantially carbon-intensive. AI's growth here depends on the pace of the region's own energy transition.

Minerals form another dimension. Indonesia's dominance of global nickel production means Southeast Asia is positioned to become central to AI not only as a site of data-center construction but as a supplier of the minerals inside the hardware itself. This layers a second, upstream environmental footprint onto the same region already absorbing the downstream infrastructure cost.

The deepest issue may be one of governance architecture rather than any single resource. ASEAN's AI governance efforts to date focus almost entirely on ethics, safety, and bias. There is essentially no regional framework covering data-center sustainability, cross-border water accounting, or lifecycle emissions reporting [41].

This leaves a structural mismatch: AI infrastructure investment is regional in scope (even global), its environmental impacts are local and often shared across borders, and the governance responding to it remains national. Closing that governance mismatch may be the more crucial step.



AI infrastructure investment is regional in scope (even global), its environmental impacts are local and often shared across borders, and the governance responding to it remains national.





What Doing It Better Looks Like

None of the organizations below have solved AI's environmental problem, and naming them here is not an endorsement of their full record. Several of the same companies profiled below for genuine engineering progress (Microsoft and Google among them) are also the companies identified earlier lobbying, via DigitalEurope, to narrow public disclosure of the very data center metrics this section relies on. The two facts do not cancel each other out; they describe two different levers (engineering versus disclosure) that a company can pull in opposite directions at once.

6.1 Companies

Google has become one of the most prominent advocates of carbon-aware computing. Rather than treating electricity as interchangeable, its systems increasingly shift computing workloads to the times and locations where low-carbon electricity is actually available on the grid, alongside hourly carbon accounting, direct renewable-energy purchasing, and machine-learning-driven cooling optimization that has measurably improved Power Usage Effectiveness across its global data center fleet. Despite this, Google's absolute electricity and water consumption have continued to rise as AI infrastructure expands, which illustrates the real distinction between improved efficiency and reduced overall impact [42].

Microsoft has arguably become the industry's most ambitious company on water specifically. Its water usage effectiveness fell from roughly **2.3** liters per kilowatt-hour in the early 2000s to about **0.27** liters per kilowatt-hour in 2025, a **90%** reduction over two decades. In June 2026, the company announced it had gone further still, claiming to have replenished more water globally than it withdrew in fiscal year 2025, hitting its 'water positive by 2030' target five years early [43].

That claim rests on replenishment projects funding watershed restoration elsewhere, not on eliminating withdrawal itself, and Microsoft's own account concedes the distinction matters: as the company put it, 'replenishing global volumes is not enough' when local watersheds can still run net-negative even as the global ledger balances. Like most of the data in this section, the underlying figures come from the company's own disclosures.

Its newest data center designs use closed-loop, chip-level liquid cooling that eliminates cooling towers and evaporative water loss entirely in many new facilities. Microsoft itself acknowledges the tradeoff: eliminating evaporative cooling increases electricity draw, since evaporation is thermodynamically cheaper than closed-loop systems.

Equinix is frequently cited by researchers for its day-to-day sustainable operations [44]:

- Strong PUE performance, improving to a global average of **1.39** in 2024 (**6%** better than the year before) with a 2030 target of **1.33**;
- Renewable electricity procurement at scale, having committed early to 100% clean and renewable energy by 2030 and reaching **96%** global coverage in 2024 [44];
- Systematic waste-heat recovery, exporting **14.5 GWh** of residual heat from its data centers in 2024, a **245%** jump from the year before;
- Water management innovation, drawing on best practices for cooling that produced a reported annual average Water Usage Effectiveness (WUE) of **0.95** in 2024, backed by customer-facing water reports and engineering solutions like water-free energy sources, thermal storage, and aggressive reuse;



- Consistent third-party sustainability reporting, offering customers granular data on electricity, renewable coverage, emissions, and water use across its **270+** data centers spanning 77 metro areas in 36 countries [45].

NVIDIA is best known for its graphics processing units (GPUs), which underpin much of today's AI infrastructure. The company's forthcoming Vera Rubin platform incorporates a direct-to-chip liquid-cooling system capable of operating with coolant temperatures of up to **45°C**, enabling facilities in favourable climates to reduce or, in some cases, eliminate the need for evaporative cooling towers [46].

NVIDIA also states that its Blackwell NVL72 platform delivers approximately **25 times** greater energy efficiency and **300 times** greater water efficiency than comparable air-cooled systems on a per-unit-of-work basis [47].

These reported efficiency gains are significant, but they should be interpreted in context. They measure the resources required to perform a given computational task rather than the total resources consumed by AI infrastructure. Blackwell NVL72 systems are designed for substantially higher computational throughput and typically operate at power levels of approximately **120–140 kW** per rack, compared with around **20 kW** for conventional enterprise server racks.

As AI deployment continues to expand, improvements in efficiency may coexist with rising aggregate electricity demand.

NVIDIA says the 45°C design can, in favorable climates, cut on-site water use by up to **100%**. These savings are true for the data center building, but they exclude the water used by the power plants generating the facility's electricity, which independent estimates put at roughly **54%** of AI's total projected water footprint through 2050 [48].

6.2 Countries

We dive into four countries for the approach they are taking.

TABLE 2.
National approaches offering transferable lessons

COUNTRY	WHY IT STANDS OUT	LESSON
Singapore	Singapore resumed data center approvals only after introducing sustainability criteria, efficiency requirements, and selective permitting, under its Green Data Centre framework. [30]	Digital growth can be conditioned on resource efficiency rather than left unrestricted.
Denmark	Denmark integrates data centers into district heating networks, recovering waste heat for homes and businesses, using heat pumps to raise the servers' low-grade cooling water to district-heating temperatures. This was made economical after Denmark abolished its surplus-heat tax for certified operators in 2022. Meta's Odense data center recovers about 215,000 MWh of waste heat a year, warming more than 12,000 homes through the Fjernvarme Fyn network and helping reduce the area's reliance on coal. [49]	AI infrastructure can contribute to broader energy systems instead of only consuming electricity.
Finland	Uses naturally cool climates, low-carbon electricity, and waste-heat recovery to cut operational impact, siting data centers next to existing district heating networks and using heat pumps to boost the low-grade heat captured from servers to distribution temperature. [50] Google's Hamina site already supplies 80% of that town's heating needs. [51] Microsoft's Espoo-Kirkkonummi cluster is expected to cover about 40% of local district heating demand once operational (roughly 100,000 homes). [52]	Geography and grid mix matter as much as technology choices.
Norway	In 2023, renewable sources provided 95% of electricity consumed in Norway; 83% came from hydropower alone, giving the grid a carbon intensity of about 15 gCO ₂ e/kWh, nearly 14 times lower than the EU average. [53]	An exceptionally clean grid can cut operational emissions to a fraction of the regional average. Though, as with any national grid, that doesn't by itself eliminate water use or embodied carbon.

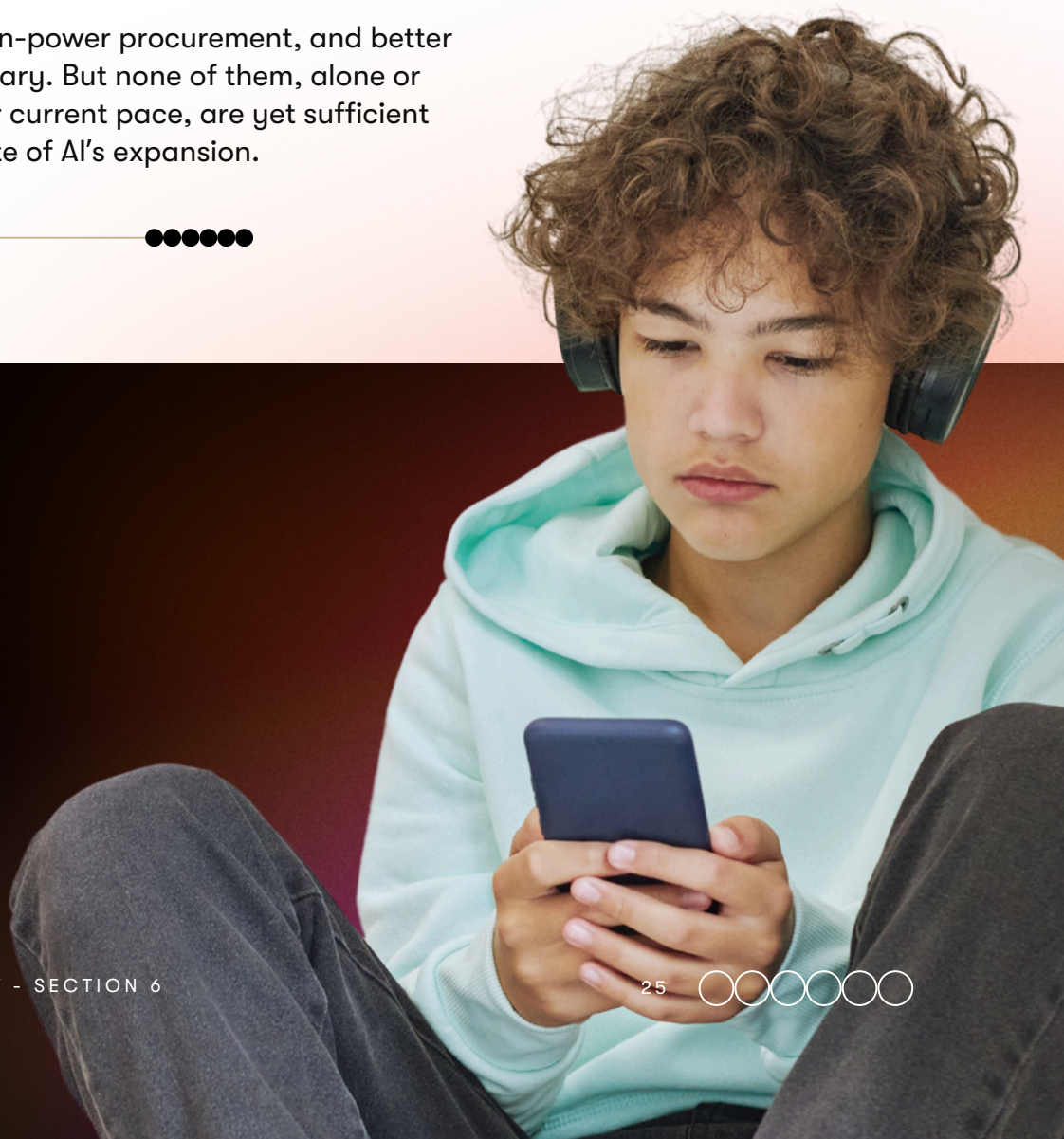


6.3. Efficiency is not the same as reduction

It would be convenient to conclude that better chips and smarter cooling will resolve this on their own. The evidence does not support that conclusion. Even as the International Energy Agency (IEA) finds that AI computing efficiency per task is improving faster than at almost any point in energy history, the same research still projects that total data center electricity demand will more than double by 2030 [7,9].

When computation gets cheaper and more efficient, people tend to run more of it, a long-observed pattern known as Jevons Paradox.

Efficiency gains, clean-power procurement, and better regulation are necessary. But none of them, alone or even together at their current pace, are yet sufficient to offset the sheer rate of AI's expansion.





What Responsible Adoption Requires

The environmental trajectory of AI is shaped by a series of decisions concerning regulation, infrastructure investment, technology design, and procurement. Responsibility therefore extends across multiple stakeholder groups, each with a distinct role in contributing to what AI's environmental footprint is going to look like.

7.1 Consumers and enterprise buyers: Driving demand for transparency

The influence of individual users over AI's environmental footprint is inherently limited. The resource demands associated with hyperscale AI infrastructure cannot be meaningfully reduced through changes in individual consumption patterns alone. Enterprise users, however, occupy a more influential position. As major purchasers of cloud computing and AI services, businesses can incorporate environmental disclosure as a condition of purchase the way they already do with security and cost.

Requiring suppliers to disclose metrics such as energy consumption, water use, carbon intensity, and renewable energy sourcing would create stronger commercial incentives for transparency and continuous environmental improvement than voluntary corporate commitments alone.



7.2 Industry: Advancing efficiency while strengthening accountability

The technology sector has demonstrated that meaningful improvements in resource efficiency are achievable. Microsoft's reported reductions in operational water intensity, Equinix's expansion of waste-heat recovery programs, and NVIDIA's continued improvements in computational efficiency illustrate that innovation can substantially reduce the environmental intensity of AI infrastructure.

However, these achievements should be viewed in context. Improvements in efficiency have not yet offset the rapid growth in overall AI deployment and the associated expansion of data-center capacity. Furthermore, while many technology companies publish sustainability commitments, public reporting on AI-specific energy use, water consumption, and lifecycle environmental impacts remains limited and inconsistent.

The next stage of responsible AI development therefore requires not only continued technological innovation, but also greater transparency, standardized reporting, and independent verification of environmental performance.

7.3 Governments: Establishing clear and enforceable regulatory frameworks

Governments play a critical role in ensuring that AI development aligns with broader sustainability objectives. Emerging policy approaches demonstrate that environmental governance need not come at the expense of digital innovation.

Singapore, for example, has linked approval for new data centers to stringent efficiency and sustainability requirements, showing that continued digital infrastructure investment can be conditioned on responsible resource use. The European Union has introduced mandatory reporting requirements for large data centers through the revised Energy Efficiency Directive while incorporating environmental transparency provisions within the AI Act.

Nevertheless, these frameworks also reveal the limitations of disclosure-based regulation when reporting requirements are weakened by exemptions or confidentiality provisions. In the

United States, state-level initiatives addressing data-center energy demand and water use illustrate that oversight can develop even in the absence of comprehensive federal legislation.

Collectively, these experiences suggest that effective governance requires regulatory frameworks that are transparent, measurable, enforceable, and sufficiently robust to prevent environmental obligations from being undermined through broad exemptions.

7.4 International cooperation: Closing the global governance gap

Perhaps the most significant challenge lies beyond national regulation. AI infrastructure increasingly operates across borders, while environmental governance remains largely organized at the national level.

The rapid development of the Singapore–Johor–Batam corridor reveals this disconnect. Digital services may be delivered from Singapore, while the electricity demand, water consumption, land use, and associated environmental pressures are increasingly concentrated in neighboring Malaysia and Indonesia. Similar cross-border dynamics are emerging in other regions as countries compete to attract AI infrastructure through access to land, energy, and favorable regulatory conditions.

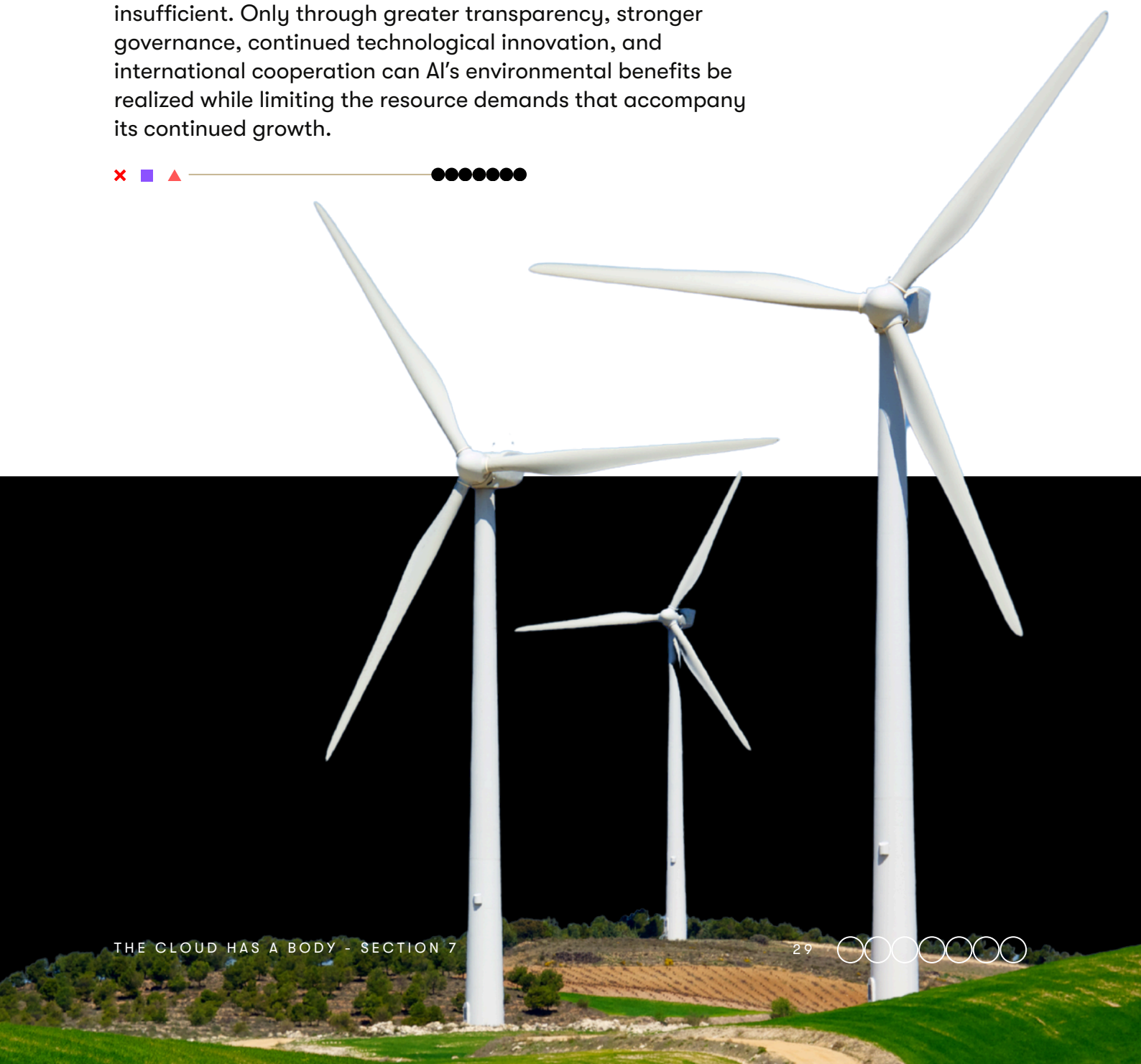
At present, no international framework systematically monitors the transboundary environmental impacts of AI infrastructure, including cross-border electricity demand, water stress, embodied emissions, or critical mineral supply chains. Existing global climate governance mechanisms primarily focus on national greenhouse gas inventories rather than the environmental footprint of digital infrastructure.

As AI deployment becomes increasingly global, greater international coordination will be required to establish common standards for environmental reporting, lifecycle assessment, and sustainable infrastructure development.

7.5 Looking ahead

The environmental consequences of AI will ultimately depend less on the capabilities of the technology itself than on the governance structures that shape its deployment. Data centers, electricity networks, cooling systems, and semiconductor supply chains represent long-term infrastructure investments whose environmental characteristics will persist for years, if not decades. Decisions taken during the current phase of AI expansion will therefore have enduring implications.

Responsible AI adoption requires coordinated action across governments, industry, enterprise users, and international institutions [54]. Progress in one area alone will likely be insufficient. Only through greater transparency, stronger governance, continued technological innovation, and international cooperation can AI's environmental benefits be realized while limiting the resource demands that accompany its continued growth.



8. References

1. Berger, E. (2026). "Sam Altman defends AI's energy toll by saying it also takes a lot to 'train a human.'" The Guardian. [Guardian] · [original remarks, TechCrunch]
2. UNEP (n.d.). "How artificial intelligence is helping tackle environmental challenges." [Link]
3. Statista (2026). "Data centers worldwide by territory 2026." [Link]
4. Programs.com (2026). "Measuring the Data Center Boom: Facts and Statistics." [Link]
5. ABI Research (2026). "How Many Data Centers Are There and Where Are They Being Built?" [Link]
6. CloudDataNexus (2026). "Exploring the Number of Data Centers by Country: A Global Overview." [Link]
7. International Energy Agency (2025). "Energy and AI." [Link]
8. International Energy Agency (2024). "Electricity 2024: Analysis and Forecast to 2026." [Link]
9. International Energy Agency (2025). "Energy Supply for AI." Energy and AI report. [Link]
10. Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). "Making AI Less Thirsty: Uncovering and Addressing the Secret Water Footprint of AI Models." arXiv:2304.03271. [Link]
11. HARC (Houston Advanced Research Center). (2026). "Thirsty Data and the Lone Star State: The Impact of Data Center Growth on Texas' Water Supply." [Link]
12. Bloomberg (2025). "The AI Boom Is Draining Water From the Areas That Need It Most." [Link]
13. UNCTAD (2024). Digital Economy Report 2024. [Link]
14. InformationWeek (n.d.). "Is AI Driving Demand for Rare Earth Elements and Other Materials?" [Link]
15. Cordillera GEO-Services (2026). "AI's Critical Minerals Problem: The Hidden Geological Cost of the Hardware That Powers It All." [Link] · [FP Analytics]
16. European Union (2024). Regulation (EU) 2024/1689 (AI Act), Article 53. [Link]
17. ScienceDirect (2026). "AI, climate, and regulation: From data centers to the AI Act." [Link]
18. European Commission (2023). Directive (EU) 2023/1791 (recast Energy Efficiency Directive), Article 12, and Delegated Regulation (EU) 2024/1364. [Link]
19. European Commission (2026). "Rating scheme for data centres in the EU — Commission launches call for feedback." [Link]
20. Normative.io (n.d.). "Scope 3 reporting under CSRD and SBTi: what you actually need." [Link]
21. Concerted Action Energy Efficiency Directive (2025). "Assessment of the energy performance and sustainability of data centres in the EU." [Link]
22. The Guardian (2026). "Microsoft, US tech firms lobbied EU for secrecy rules on datacentre emissions." [Link]
23. The White House (2025). Executive Order 14110 (rescission record). [Link]
24. The White House (2025). Executive Order 14318, "Accelerating Federal Permitting of Data Center Infrastructure." [Link]
25. NIST (2023). AI Risk Management Framework (AI RMF 1.0), NIST AI 100-1. [Link] · [AI Risks and Trustworthiness]
26. Duane Morris (2026). "SEC Proposes Rescission of Climate-Related Disclosure Rules." [Link]



27. MultiState (2026). "State Data Center Policy 101: 2026 Legislative Guide." [Link]
28. ArentFox Schiff (2026). "State Regulation of Data Centers in 2026 — A Shifting Landscape." [Link]
29. Climate XChange (2026). "State Policy Toolkit for Data Center Regulation on Water Impacts." [Link]
30. Infocomm Media Development Authority of Singapore & Economic Development Board (2024). Singapore Green Data Centre Roadmap. [Link]
31. Knight Frank (2025). Southeast Asia data centre market research — Johor operational capacity estimates. [Link]
32. Cushman & Wakefield (n.d.). Analysis of hyperscaler site-selection drivers in Johor, Malaysia. [Link]
33. BusinessToday Malaysia (2026). "Cloud Seeding To Tackle Water Shortage In Kota Tinggi As 4,000 Residents Affected." [Link]
34. PUB, Singapore's National Water Agency (n.d.). "Imported Water." [Link]
35. Malaysia National Water Services Commission, reporting via The Independent Singapore (n.d.). "Johor's data centre boom raises fears over mounting water strain." [Link]
36. DataCenterDynamics (n.d.). "Indonesia's next strategic AI and data center gateway." [Link]
37. ANTARA News (n.d.). "BP Batam oversees US\$5 billion investment in AI data center." [Link]
38. Intimedia (n.d.). "Massive data center investments worth hundreds of trillions flow into Batam." [Link]
39. World Energy Council Malaysia (2025). "World Energy Trilemma: Malaysia." [Link]
40. Katadata (2025). "Indonesia's Electricity Generation Energy Mix 2015–2024." [Link]
41. CSIS (n.d.). "Beyond the Matrix: AI Governance Gaps in Southeast Asia." [Link]
42. The Guardian (2025). "Call to make tech firms report data centre energy use as AI booms." [Link]
43. Windows News. (2026). "Microsoft's Datacenter Water Efficiency Improves 88% to 0.27 L/kWh as AI Scales." [Link]
44. Equinix Newsroom (2025). "Equinix Named a Leader in the IDC MarketScape: Worldwide Datacenter Services Sustainability 2025–2026 Vendor Assessment." [Link]
45. Equinix Newsroom (2025). "Equinix to invest £3.9 billion in UK's critical national infrastructure." [Link]
46. Cooling Post (2026). "NVIDIA raises data centre cooling to 45oC." (Vera Rubin platform.) [Link]
47. NVIDIA (2025). "Blackwell Platform Boosts Water Efficiency by Over 300x." [Link]
48. Tech Times (2026). "AI Data Center Water Use Is Not Solved: Nvidia's Cooling Fix Stops at the Walls." [Link]
49. Ramboll (n.d.). "Meta: Surplus Heat to District Heating." [Link]
50. The Northern Voices (2026). "Turning Data Centers Into City Heaters: Finland's Answer to the Environmental Cost of the Cloud." [Link]
51. Google (2024). "Our first offsite heat recovery project lands in Finland." [Link]
52. Global Brands Magazine (n.d.). "Microsoft Data Centre in Finland." [Link]
53. stw.no (n.d.). "Norway: A High-Growth Data Centre Market Built on Hydropower." [Link]
54. UNEP (2025). "AI has an environmental problem. Here's what the world can do about that." [Link]

Disclaimer

This publication is produced by Effective Thoughts for general informational and educational purposes only. It does not constitute professional, legal, regulatory, or investment advice.

The content draws on publicly available sources believed to be reliable at the time of writing. While reasonable care has been taken in its preparation, Effective Thoughts and its collaborating partners have not independently verified all underlying data and make no representation or warranty, express or implied, as to its accuracy, completeness, or currency. Information may change after the date of publication.

References to specific organizations, technologies, jurisdictions, or projects do not imply endorsement. The views expressed are those of the authors and do not necessarily reflect the official positions of the collaborating organizations, their staff, partners, or clients.

Effective Thoughts and its collaborating partners accept no liability for any loss or damage arising from reliance on the contents of this publication.

© 2026 Effective Thoughts. Licensed under CC BY-NC-ND 4.0. You may share this publication in full for non-commercial purposes with attribution; no modified versions may be distributed.