2026
February

Sustainability Implications of Rising Energy Demand from Data Centres and Generative AI

Sustainability Research Paper

The Al-Attiyah Foundation

01 OUR MEMBERS

The Al-Attiyah Foundation is proudly supported by:

Global data centre electricity demand has roughly doubled since 2020 and the International Energy Agency (IEA) has projected it will more than double again by 2030, driven primarily by generative Artificial Intelligence (AI) model-building and deployment activities. This surge has prompted capital commitments on an unprecedented scale, including financing by sovereign wealth funds across the GCC. However, investment is outpacing the grid, water and planning systems on which the AI facilities will depend.

This paper examines the sustainability-related risks and opportunities that stem from this imbalance. It analyses physical climate exposure, financial materiality and technology risk, the regulatory response across the EU, the U.S. and the MENA region, and evaluates emerging strategies—from behind-the-meter power (BTM) and waste-heat products to green finance instruments.

SUSTAINABILITY RESEARCH PAPER

This research paper is part of a 12-month series published by the Al-Attiyah Foundation every year. Each in-depth research paper focuses on a current sustainability topic that is of interest to the Foundation's members and partners. The 12 technical papers are distributed to members, partners, and universities, as well as made available on the Foundation's website.

- Global data centre electricity demand has roughly doubled since 2020. The IEA has projected that it will more than double again by 2030, driven primarily by generative AI model-building and deployment activities that require significantly more power than conventional cloud computing.
- The sustainability risks associated with data centres are increasing. Over half of the world's top 100 data centre hubs will face high or very high-water stress by 2030, with the Middle East facing the highest exposure, compounding the physical climate risk from Cooling Degree Days (CDD) and grid-reliability incidents already observed in places such as Virginia, London and the Iberian Peninsula.
- Against this backdrop, financial materiality has grown rapidly. Capital requirements for AI-ready data centres have attracted sovereign and institutional commitments on an unprecedented scale, but the cost structure—driven by energy procurement, water, hardware obsolescence and rising interest rates—has introduced material risks alongside strong revenue prospects for data centre projects.
- Social and system-level impacts are becoming apparent. Large-scale grid "departures" by data centre operators have exposed vulnerabilities in electricity networks, while competition for water resources and community opposition have blocked or delayed significant planned investment in both Europe and North America.

- Management and governance frameworks are tightening at the same time. The EU's Energy Efficiency Directive (EED) now requires annual reporting against 24 key performance indicators for data centres above 500kW, yet implementation has been hampered by non-cooperation by tenants, confidentiality disputes and measurement inconsistencies.
- Finance and risk transfer mechanisms remain underdeveloped for data centres. Traditional insurance leaves significant residual risk for physical and business-interruption losses. Parametric products, which pay out automatically when a pre-defined trigger (such as a temperature threshold) is met, might offer faster, more transparent coverage if they become established.
- "Adaptation and resilience" investment has become critical for data centre operations. High-efficiency cooling, alternative water sources, resilient backup power and waste-heat reuse are becoming essential requirements for facilities in heat- or water-stressed regions.
- Technology risk is accelerating, while regulatory frameworks have not kept pace. Hardware lifecycles of 18–24 months conflict with the longer depreciation schedules, creating stranded-asset exposure. Meanwhile, the EU AI Act, the EED and national grid-connection policies are setting new disclosure benchmarks.
- Corporate strategy favours vertical integration for data centre operation. Leading operators and "hyperscalers" have pursued behind-the-meter (BTM) power generation, co-located renewables and circular hardware models to reduce grid dependency.
- GCC states have emerged as significant AI-energy hubs, with regional data centre capacity expected to triple by 2030, driven by sovereign AI projects in the UAE and Saudi Arabia. Chronic water scarcity and extreme heat make cooling design especially challenging.
- Yet the opportunity is substantial. Green asset-backed securities for data centres have grown rapidly, waste-heat monetisation has reached commercial scale and AI-enabled digital twin platforms have demonstrated measurable efficiency gains.

05 DATA CENTRES AS A SOURCE OF COMPLEX AND CORRELATED RISKS

The growth in demand for computing, driven by AI, has pushed data centre electricity consumption well beyond the pace at which low-carbon supply can feasibly be provided. The IEA projects that natural gas will remain the principal source of incremental generation for data centres through 2030, with the U.S. accounting for the largest share of growth. Renewables, though supplying approximately half of present requirements, cannot yet match the constant "load profile" that such facilities demand without more dependable reserve capacity.

Current dependence on gas introduces a transition risk of a kind that is not easily hedged. National regulators have begun imposing renewable-procurement requirements on data centre operators, yet long transitional periods will sustain gas-fired generation well into the next decade. Where grid interconnection queues discourage reliance on renewable procurement, operators have resorted to behind-the-meter gas and diesel plants, adding to emissions and creating balance-sheet exposure to increasingly stringent carbon-pricing regimes.[1,2]

At data centre sites, heat and water stress have moved from the category of theoretical hazard to a cause of actual operational failure. In 2022, cooling failures at two London data centres serving Guy's and St Thomas' during record temperatures exceeding 40 °C caused service outages lasting six weeks and direct costs of approximately $1.7 million. Climate risk analysis of the hundred leading global data centre locations indicates that roughly three-quarters will face materially higher cooling demand, with a substantial rise in CDD between 2030 and 2080. More than half of the present stock of data centres already confronts high or very high-water stress, a proportion projected to increase further by 2050.[3]

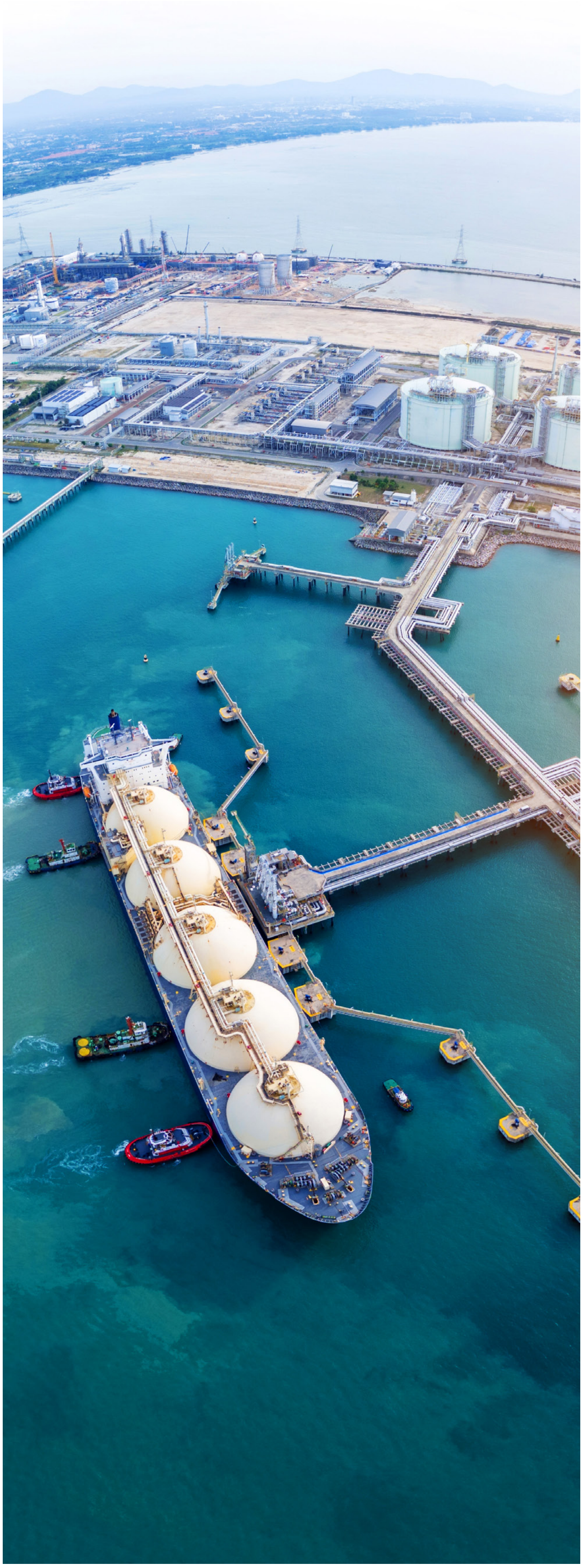
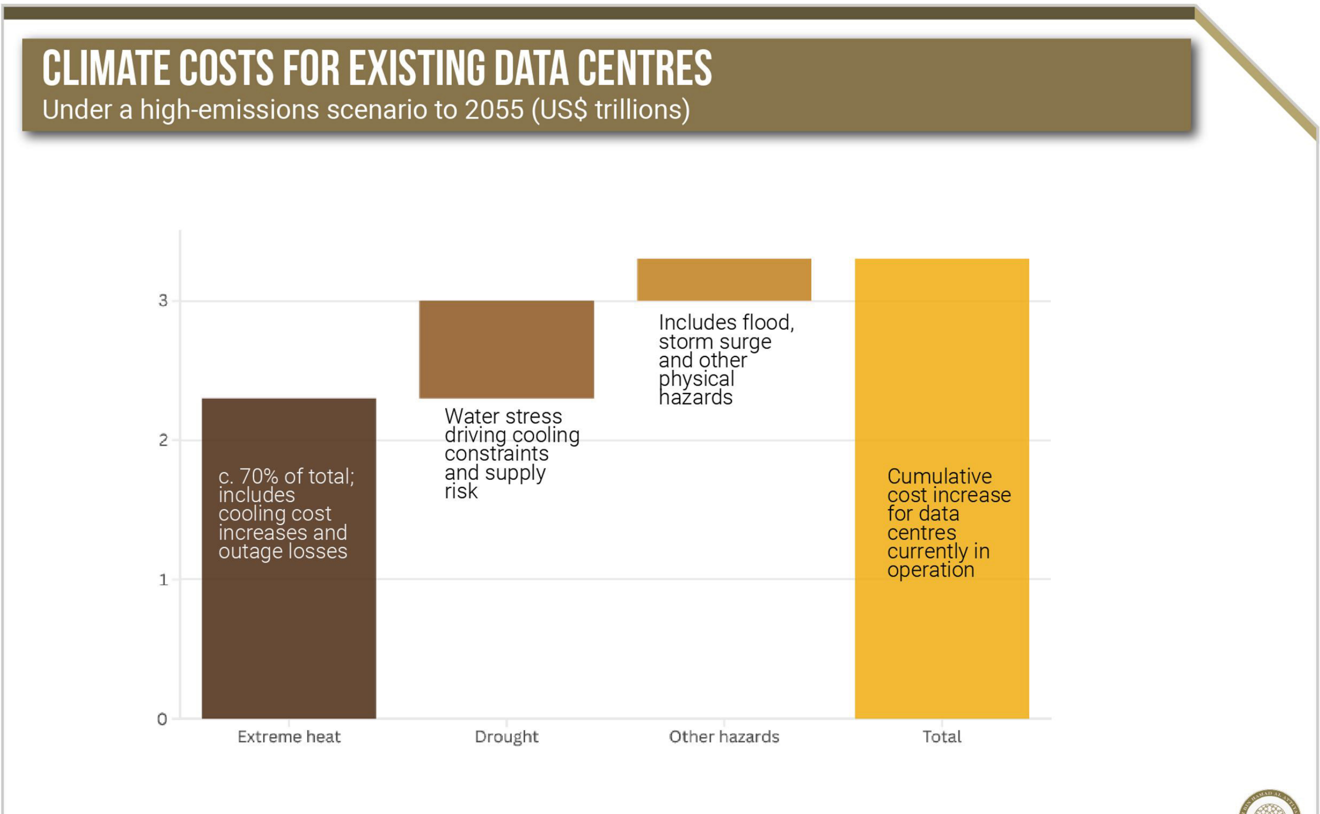


Figure 1: Cumulative Climate-Driven Costs for Existing Data Centres to 2055

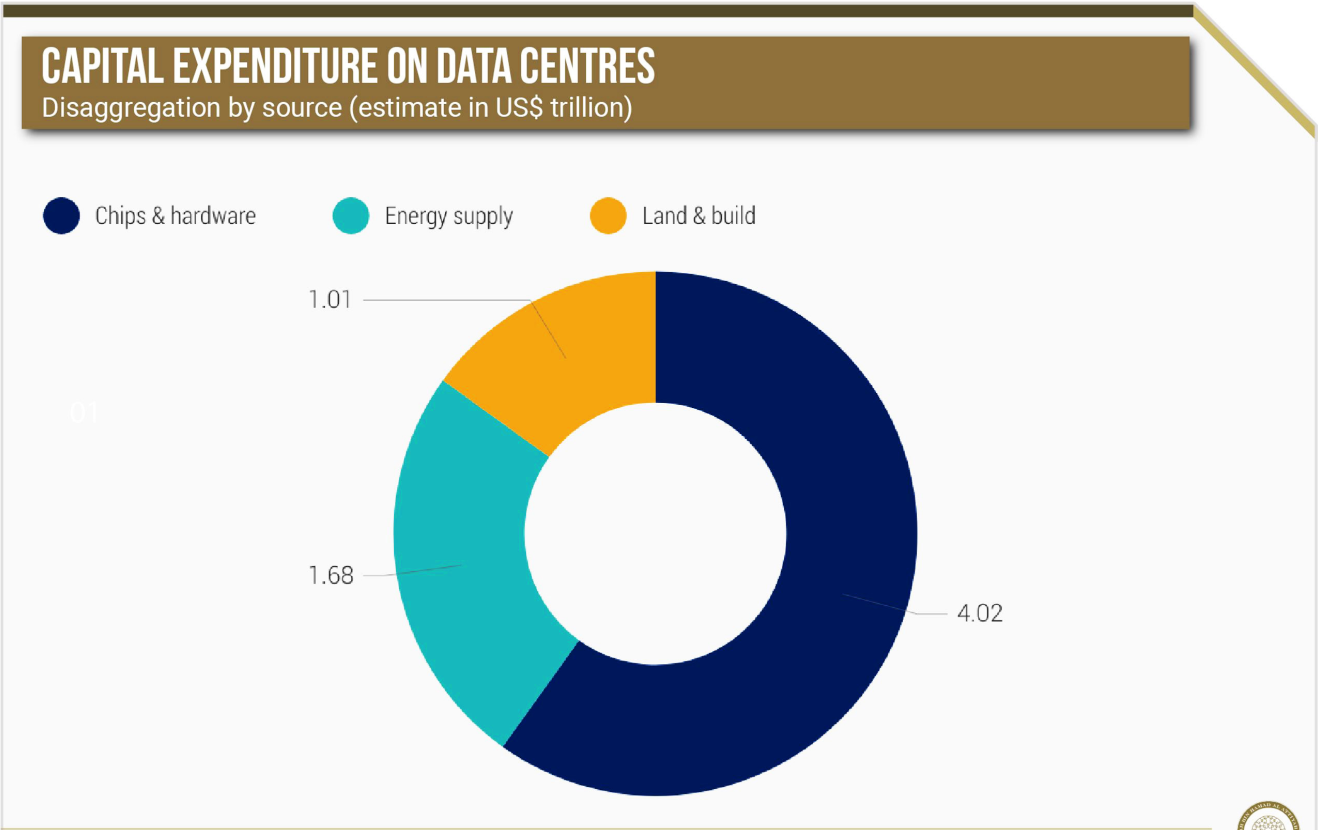


Einhorn, G. and King, D. (2025) 'The $3.3 trillion climate question: can data centres take the heat?', World Economic Forum, 10 October. Based on S&P data.

The financial consequence is highly material. The World Economic Forum (WEF), drawing upon S&P data, estimates that cumulative climate-driven cost increases for data centres presently in operation could reach $3.3 trillion by 2055 under a high-emissions pathway—equivalent to nearly 10% of total asset value—with extreme heat accounting for more than two-thirds of the impact and drought for much of the remainder. Without targeted structural adaptation, insurance premiums are likely to rise, and a growing share of facilities will pass into high-risk classifications.[3,4]

The insurance market has not kept pace. Conventional property and business-interruption policies assume ordinary commercial premises and do not accommodate the cascading failures specific to data centres, where a cooling malfunction can destroy server equipment within minutes and expose the operator to simultaneous liability claims from multiple tenants. Capacity is concentrated in a small number of locations—London, northern Virginia, Dublin, Singapore—and the resulting "accumulation risk" limits underwriting appetite. Specialist underwriters able to evaluate liquid cooling, cyber-physical issues and AI-driven building management remain scarce. Parametric instruments, which settle automatically against a pre-defined trigger such as a temperature threshold, could offer faster and more transparent coverage, yet adoption is limited by the absence of standardised "loss indices" for data centres and by uncertainty over how parametric pay-outs interact with indemnity contracts.

Figure 2: Data Centre Capital Expenditure Allocation to 2030



Noffsinger et al (2025) 'The cost of compute: a $7 trillion race to scale data centers', McKinsey Quarterly, 27 April.

Until these mechanisms mature, operators retain a substantial share of climate-related loss on the balance sheet.

Hardware obsolescence is another factor to consider when assessing exposure. New GPU generations now appear every 18 to 24 months, yet depreciation schedules commonly extend to five or six years; the technology is therefore rendered commercially obsolete well before its cost has been written off. This mismatch interacts with physical climate risk and transition pressure to produce a compound exposure. Extreme heat raises cooling costs, strains grids and shortens the useful life of equipment. Grid instability during heatwaves forces reliance upon diesel-based backup sources, compounding emissions and regulatory risk.

For operators and investors, the central governance challenge is that physical, transition and technology risks are correlated. The situation demands integrated scenario analysis rather than treating each hazard in isolation.

Data centres have become one of the largest infrastructure asset classes in the world in less than a decade. McKinsey estimates that global data centre capital expenditure could reach $6.7 trillion by 2030, of which AI-specific facilities would account for $5.2 trillion and traditional IT infrastructure for the residual $1.5 trillion. With approximately 70% of new demand driven by artificial intelligence, global data centre capacity could almost triple to roughly 219 gigawatts (GW).

There is a considerable element of uncertainty inherent in forecasts for so new an asset class—McKinsey's scenario range for total investment spans $3.7 to $7.9 trillion—yet even the conservative lower bound exceeds projected global upstream oil and gas exploration expenditure over the same horizon.[5]

Roughly 60% of data centre investment is expected to be committed by technology developers for chips and computing hardware, 25% by energy suppliers for generation, transmission and cooling and 15% by builders for land and site development. The concentration of capital expenditure in hardware points to an ever-closer dependence of data centre asset-class performance upon the product cycle of semiconductor manufacturers and upon global chip supply-chain constraints.

Demand for AI model building and deployment has driven rapid revenue growth in data centre markets. "Hyperscale" operators—providers that manage vast server facilities at a global scale—now account for a substantial share of earning capacity, while "Neoclouds"—a newer cohort of specialist operators constructed to host GPU-intensive AI setups—have emerged alongside them. Cost pressures are nevertheless evident. Energy has become the single largest operating expense in high-density facilities. Moreover, GPU rental rates are falling as supply outpaces demand—a particular hazard for debt-financed Neoclouds—and unplanned downtime costs can run into the thousands of dollars per minute. The financing architecture for AI data centres has diversified rapidly. Brookfield Asset Management launched a $100 billion AI infrastructure programme in November 2025, with NVIDIA and the Kuwait Investment Authority as founding partners. The Qatar Investment Authority committed $20 billion through its Qai joint venture with Brookfield in December 2025.

There are emerging indications that the binding constraint on data centre growth is no longer financing or demand, but implementation. Power availability, regulatory friction and grid delays—up to seven years in the U.S. for facilities exceeding 100 megawatts—are preventing projects from reaching the construction phase. In December 2025, Fermi America—co-founded by former U.S. Energy Secretary Rick Perry—lost between 30% and 50% of its market capitalisation after a tenant cancelled a $150 million construction agreement for its Texas-based AI campus. The case illustrates how counterparty, technology and project implementation risks interact across the data centre value chain.

09 SOCIAL AND SYSTEM IMPACTS

A succession of major disruptions to critical services in 2025 demonstrated how failures of cloud infrastructure can propagate across entire economies. In April 2025, a complex grid failure disconnected 31 GW of generation capacity across Spain and Portugal, shutting down payment services, airports, telecommunications and internet access for up to ten hours and affecting over 60 million people. In October, successive outages at Amazon Web Services and Microsoft Azure cascaded across enterprise and consumer platforms. In November 2025, a Cloudflare disruption affected roughly a fifth of global web traffic.[6]

Grid infrastructure is itself a source of systemic risk. In July 2024, a single component failure on a high-voltage transmission line in Virginia, U.S., forced 60 data centres consuming 1,500 MW to disconnect simultaneously and switch to backup generators. Energy firms were required to curtail voltage across the network to avert regional blackouts. As AI workloads concentrate in fewer, larger facilities, the exposure of critical digital services—healthcare records, government benefits and financial payment operations—to a single point of failure increases materially.[7]

Data centre cooling places significant and contested claims upon freshwater resources. Planning authorities in Australia approved ten data centres in Sydney without requiring measurable water-reduction plans, notwithstanding projected annual consumption of roughly 2% of the city's maximum supply. Projections indicate that data centres could account for up to a quarter of Sydney's available water by 2035, assuming operators meet cooling efficiency targets.[8]

Transparency is likely to be a persistent source of contention. In late 2025, California vetoed a bill that would have required data centres to disclose annual water consumption and anticipated usage to their water suppliers when applying for or renewing licences. The veto cited concerns about business impact amid "unprecedented demand" for data centres driven by AI adoption. A medium-sized facility can consume up to 110 million gallons annually—equivalent to the domestic supply for roughly 1,000 households—yet few jurisdictions require site-level public disclosure.[9]

Resistance to data centre siting has become a material development risk. Monitoring bodies covering the proposed sites of major new constructions have attested to a surge in opposition, with some projects blocked or delayed amid local resistance, affecting billions of dollars in potential investment. The disruption is pronounced in the US. In Kansas, petitions and lawsuits have paused a $12 billion hyperscale campus. In Pennsylvania, residents drafted their own restrictive zoning legislation. Georgia and Indiana have begun formalising regulatory reviews for large facilities.[10]

The reduction of tax abatements has compounded the difficulty in the US. In Minnesota, a sales tax exemption on electricity for data centres—previously worth as much as $100 million per year to data centre operators—prompted Amazon to suspend a planned data centre. For sponsors of large-scale projects, including Gulf-backed ventures seeking US sites, community consent and regulatory predictability have become as important as power availability and land cost.

Leading data centre operators have begun embedding climate, energy and technology risks into formal enterprise risk management (ERM) frameworks for scenario analysis.

Kyndryl, for instance, incorporated 91 data centre sites in its 2025 physical risk assessment, linking short- and long-term climate scenarios to financial planning and governance responses. This practice remains the exception rather than the norm, though, as many operators—particularly smaller colocation providers—have yet to formalise data centre physical risk within board-level risk registers.[11]

The operational backbone for these frameworks typically rests upon ISO 14001 (environmental management) and ISO 50001 (energy management). Certification signals process maturity but does not, of itself, guarantee that climate risk is integrated into capital allocation or strategic planning. For infrastructure investors, the governance question is whether operators treat climate and energy as enterprise-level risks subject to board oversight or as compliance tasks delegated to operational teams.

Boards and senior technology leaders face growing pressure to assume accountability for energy strategy, resilience and emissions. A Beyond Fossil Fuels poll in October 2025 found that a majority of European citizens support binding rules to limit data centres' impacts on energy, water and land use, adding reputational risk for operators that fail to demonstrate governance. Several hyperscale operators have integrated sustainability indicators—including power usage effectiveness (PUE), water usage effectiveness (WUE), and renewable energy procurement—into existing commercial risk reporting at the board level. Yet this integration has not necessarily improved targeted outcomes: the Uptime Institute's latest global data centre survey found that average PUE has been substantially unchanged for many years, constrained by legacy infrastructure and climate-related barriers to cooling.[12]

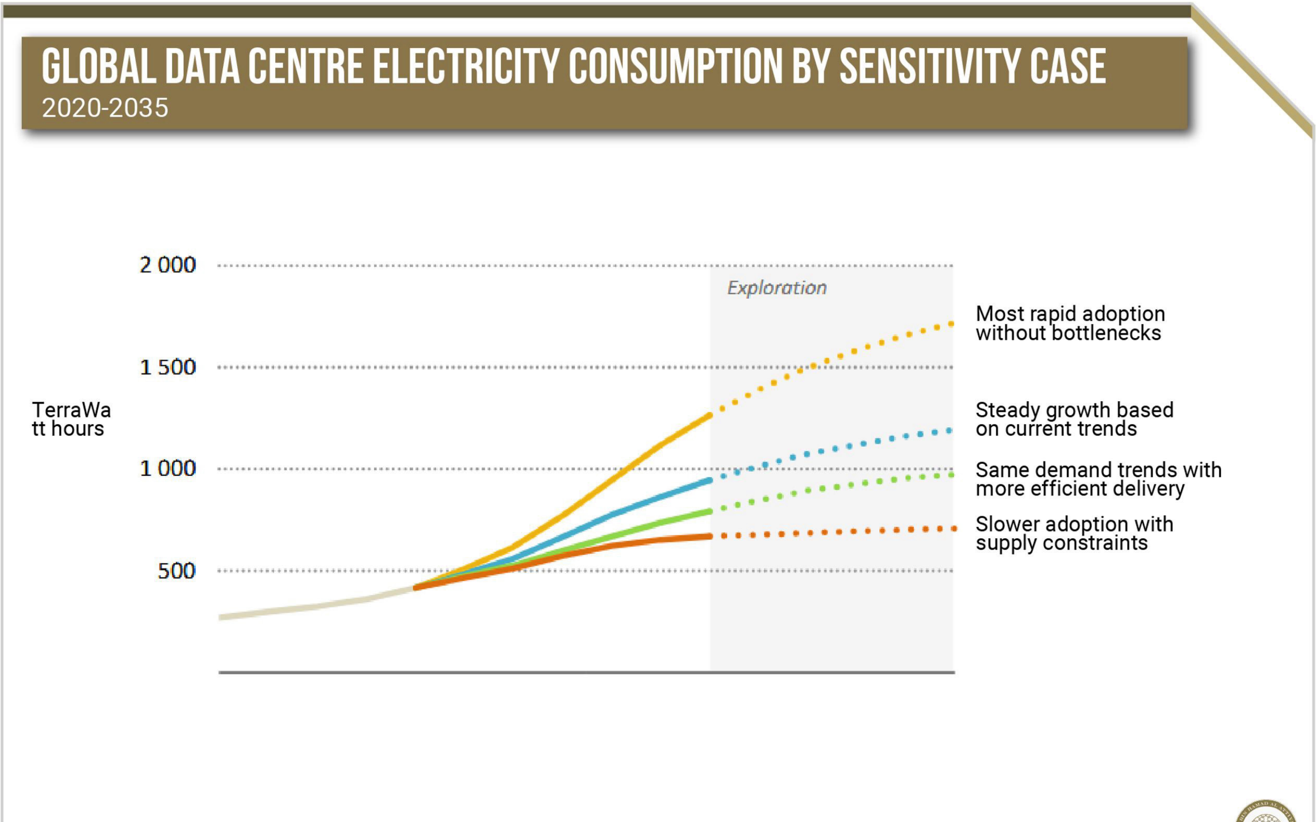
The European Commission has become the dominant regulatory force in data centre sustainability governance13. The EU AI Act requires providers of general-purpose AI (GPAI) models to document and disclose the energy consumed during training. Models whose training compute exceeds a defined threshold are assumed to carry "systemic risk" and face mandatory risk assessment, cybersecurity and incident-reporting obligations. The supporting Code of Practice reinforces these requirements by mandating ongoing documentation of the model, including energy use.14 Europe's leading role in legislating for data centres is in stark contrast to the lagging international role of its corporates and institutions in terms of data centre investment and overall sector growth. Beyond the reporting regime itself, AI model opacity compounds the governance challenge. In multi-tenant environments, firms often cannot measure or audit energy consumption on a per-query or per-workload basis. This limits their ability to set meaningful efficiency targets or allocate capital to the highest-impact improvements.

Moreover, the rapid adoption of AI in cyber-physical systems has reshaped the data centre risk profile, with industry analysts warning that misconfigured AI embedded in national infrastructure could disrupt critical services in major economies. The UK government's Cyber Security and Resilience Bill, introduced in November 2025, has responded by designating data centres as operators of essential services, with colocation facilities above 1MW and enterprise facilities above 10MW subject to mandatory 24-hour incident reporting.[15]
The overall context strengthens the case for adaptation investments to mitigate physical risks. A detailed study, analysing nearly 9,000 data centres globally, found that targeted structural measures—mainly cooling upgrades and flood protection—could reduce the number of high-risk facilities by more than two-thirds by 2050.

Without intervention, climate hazards threaten over a quarter of the global stock of data centres. For operators in heat-stressed regions, the implication is that cooling design is no longer an efficiency choice but a resilience investment with direct bearing upon asset "insurability".[4]

"Liquid cooling"—whereby coolant is delivered directly to processors rather than circulated as chilled air—has become a mainstream feature of data centre facility design. Industry sources report that direct-to-chip systems can reduce cooling energy and water consumption by up to 60% while supporting rack densities above 100kW. The broader data centre liquid cooling market reached an estimated $5.5 billion in 2025, with direct-to-chip solutions accounting for around 40% and showing the fastest growth. For facilities hosting high-density AI workloads, air-based cooling alone may no longer be viable.

Figure 3: Data Centre Electricity Demand Forecasts



IEA (2025), Energy and AI, IEA, Paris https://www.iea.org/reports/energy-and-ai, Licence: CC BY 4.0.

Proactive "load shifting" depends upon integrating climate forecasts into operational decision-making. The World Meteorological Organization (WMO) published guidance in December 2025 on early warning systems for the energy sector, linking meteorological monitoring to grid and facility-level resilience planning. Several operators have begun applying Intergovernmental Panel on Climate Change (IPCC) physical scenarios across short- and long-term horizons to highlight risk and guide load-management decisions.[16]

Diesel generators remain the default backup for most facilities, but hydrogen fuel cells have been deployed in pilots. A 3 MW hydrogen module launched in September 2025—the highest-power stationary fuel cell unit yet documented—can offer up to 48 hours of zero-emission backup power and be transported in a single shipping container. Meanwhile, proton exchange membrane (PEM) fuel cells can reach full power within minutes and are tolerant to frequent "start-stop" cycles, which makes them well-suited to AI workloads where an uninterruptible supply is critical.

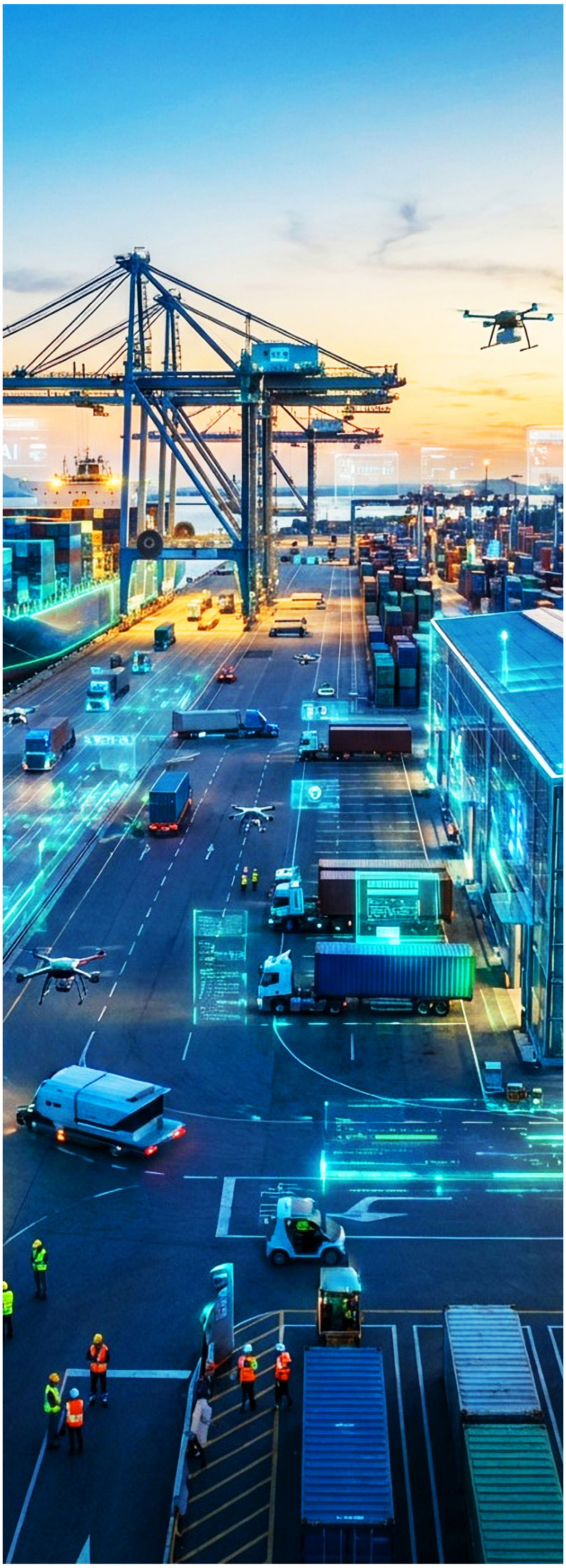
Waste-heat capture has continued to grow in Europe. As an example, a new Danish data centre announced in November 2025 is expected to supply enough recovered heat for approximately 6,000 homes via Copenhagen's district heating network. This innovative project demonstrates that operators can convert what might otherwise be regarded as a thermal liability into a source of infrastructure revenue at commercial scale.

Peer-reviewed research published in Nature Cities has demonstrated that algorithmic efficiency gains in metropolitan data centres have counterintuitively added to, rather than reduced, the energy footprint of generative AI platforms. The analysis is based on evidence from facilities concentrated in major urban technology hubs. The findings offer corroboration of the so-called "Jevons paradox"—the phenomenon whereby per-unit efficiency improvements stimulate additional demand sufficient to increase overall consumption.[17]

A systematic review of 150 articles published in Frontiers in Energy Research confirmed that AI-driven productivity gains reduce energy per computation yet frequently lead to higher total energy use, from a study of observations across data centres, transport and manufacturing. The aggregate impact may actually eliminate the environmental benefit of energy efficiency gains.[18]

Average power consumption per server rack has continued to rise, with most new installations now operating in the 10-30 kW range. This range represents a case of extreme density relative to five years earlier. A combined Vertiv-NVIDIA study documented a 10% reduction in total data centre power when moving to direct-to-chip liquid cooling, with a separate peer-reviewed study in the ASME Journal of Electronic Packaging (2024) recording a comparable 12% power reduction. Although the shift to liquid cooling is accelerating in new builds, retrofitting remains capital-intensive and complex.[19]

Hardware turnover at this pace generates increasing volumes of electronic waste. The 18-to-24-month GPU replacement cycle means that data centres now contribute a significant and rising share of global e-waste, yet the sector lacks dedicated recycling infrastructure.



Only a fraction of decommissioned data centre equipment is currently recycled through certified processes. Microsoft's Circular Centers programme—the most advanced initiative publicly documented—achieved a 91% server reuse and recycling rate in 2024 by designing for disassembly and recovering components at scale. Few other operators have matched this standard. The absence of sector-wide take-back obligations or extended producer responsibility frameworks for AI-grade hardware means that precious metals, rare-earth elements and high-value components are frequently sent to landfill rather than returned to the supply chain. For regulators preparing data centre sustainability standards, circularity metrics represent an obvious gap in the current disclosure architecture.[20]

Converging Compliance Regimes

The European Commission's data centre reporting regime under the recast Energy Efficiency Directive (EED) has moved into its second cycle, but implementation has been troubled. The Uptime Institute reported that several hyperscaler and enterprise tenants refused to share data with colocation hosts and that the first year's reported data was largely unusable for policy purposes; most EU member states had only partially transposed the relevant provisions of Article 12 by early 2025.[12]

The Commission has signalled a significant acceleration and seeks to adopt a Data Centre Energy Efficiency Package (DCEEP) in Q1 2026. This will encompass a delegated regulation on a common sustainability rating scheme and a call for evidence on binding minimum performance standards (MPS) covering power usage effectiveness (PUE), water usage effectiveness (WUE) and renewable energy factor (REF).[21]

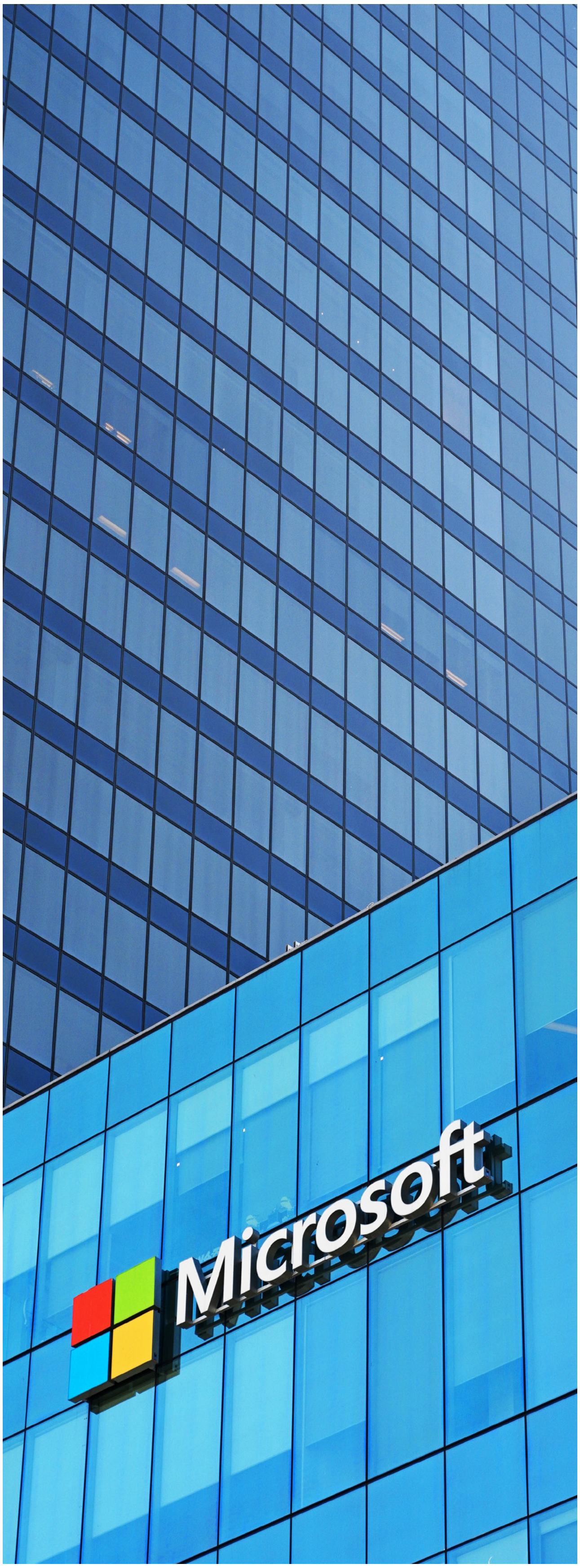
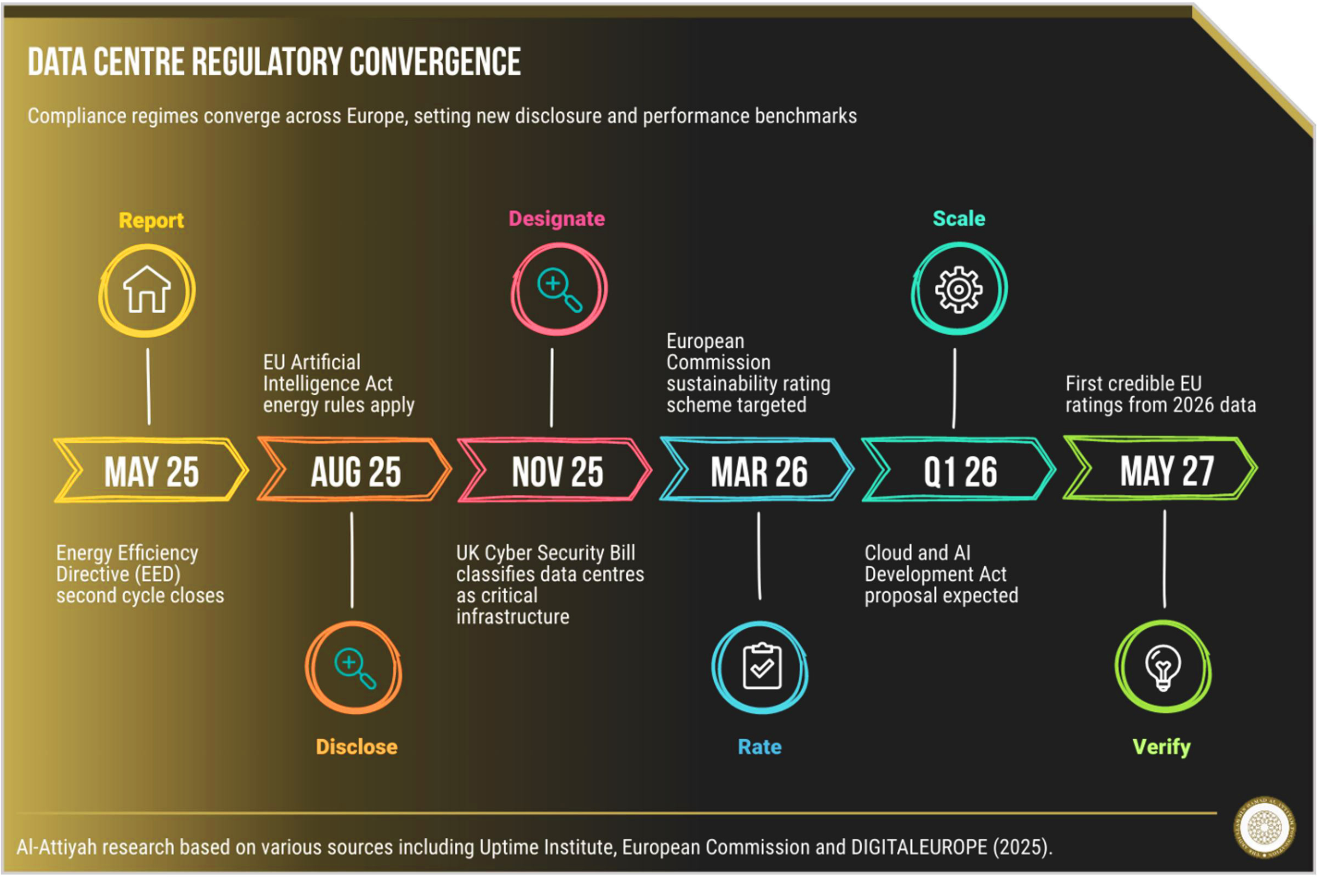


Figure 4: Data Centre Regulatory Convergence Timeline (2025-2027)



Commentators have argued that the earliest credible date for a first rating publication is mid-2027, based on calendar year 2026 data, and have cautioned against including Scope 1/2 greenhouse gas emissions or cooling technology type in the reporting criteria.[22]

Ireland's Commission for Regulation of Utilities (CRU) published a new electricity connection policy in December 2025, ending a four-year de facto moratorium on new data centre connections. Operators must now meet 80% of annual electricity demand with additional renewable electricity generated in Ireland, with a six-year compliance path. Large facilities must also provide reliable on-site or nearby generation and storage, participating in the wholesale electricity market.[23]

The International Sustainability Standards Board (ISSB) IFRS S2 Climate-related Disclosures standard—which integrated and built upon the former Task Force on Climate-related Financial Disclosures (TCFD) framework—has continued to raise expectations for quantified scenario analysis and financial impact estimation of both physical and transition risks. The IFRS Foundation published new guidance in June 2025 on transition plan disclosure, building on the Transition Plan Taskforce (TPT) framework, which the Foundation assumed responsibility for in 2024.[24]

The gap between expectations and practice has remained wide. A survey by the EDHEC Infrastructure and Private Assets Research Institute found that 66% of infrastructure investors had not evaluated physical climate risk themselves, despite 97% acknowledging such risk as significant—a finding of direct relevance given data centres' fixed-location, long-lived asset profiles and exposure to heat, water and grid hazards.[25]

The European Commission has been preparing the Cloud and AI Development Act (CADA), with a legislative proposal expected in the first quarter of 2026. The initiative has aimed to treble EU data centre capacity within five to seven years, address permitting, energy, water and land access barriers and define requirements for highly secure EU-based cloud and AI computing capacity.21
The contest for reliable baseload power has become the defining strategic variable for data centre operators.

The Federal Energy Regulatory Commission (FERC) issued a unanimous order on 18 December 2025 directing PJM Interconnection—the grid operator serving 13 U.S. states—to reform its tariff rules for generators co-located with large loads, finding the existing rules "unjust and unreasonable." The order established three new transmission service options—firm contract demand, non-firm contract demand and interim non-firm—designed to compress interconnection timelines by reducing the scope of required transmission upgrades.[26]

Behind-the-meter (BTM) power generation has been codified as a distinct asset class. A $5 billion framework agreement between Brookfield and Bloom Energy, announced in October 2025, targets up to 1 GW of fuel-cell capacity deployed directly at AI data centres, bypassing grid constraints entirely.

The IEA has projected that natural gas-fired generation for data centres will grow by 175 TWh to 2030, with much of this growth accruing through BTM measures in the US. Renewables—primarily wind, solar photovoltaic and hydro—accounted for 27% of data centre electricity in 2024 and are expected to reach approximately 50% by 2030. Still, their intermittency has remained incompatible with the round-the-clock "load profile" that AI workloads demand.[1]

Site selection criteria have shifted materially. Power availability, water access and permitting speed have displaced traditional connectivity and land-cost metrics as the primary determinants. Oxford Economics has forecast that UK data centre electricity demand alone could grow fivefold over the next five years, illustrating the grid pressure that compels operators to seek locations with both grid headroom and behind-the-meter optionality.[27]

Water consumption across the global stock of data centre facilities is projected to double by 2030. Saving water through air cooling requires significant electricity, creating a trade-off that demands integrated water-energy planning. Water purchase agreements (WPAs)—long-term frameworks modelled on power purchase agreements (PPAs)—are emerging as a mechanism to secure water supply and share infrastructure investment with local water utilities. No standardised WPA template has yet been established, but the similarity to PPA structures may accelerate adoption as water stress intensifies.

Multi-fuel strategies have gained popularity as operators seek to eliminate the risk of a single point of failure in energy supply. A gas-to-nuclear hybrid model has been developed in which natural gas generation provides an immediate source of "bridge" power while small modular reactors (SMRs) are expected to be introduced for commercial operation from the late 2020s. Significant SMR capacity is in active development pipelines for data centre co-location, with large commitments from major hyperscalers.[28]

Table 1: Gulf Cooperation Council Sovereign AI-data Centre Commitments

GULF STATES COMMIT OVER $54 BILLION TO SOVEREIGN AI DATA CENTRES

Selected GCC data centre investment commitments in 2025

	Details
Qatar ($21bn)	Qai joint venture with Brookfield ($20bn) plus Syntys sovereign AI cloud ($1bn)
Kuwait ($20bn)	Brookfield AI Infrastructure Fund with NVIDIA and Kuwait Investment Authority
UAE ($10bn)	Stargate UAE via G42, OpenAI, Oracle, NVIDIA and SoftBank; targets 5 GW total capacity
Saudi Arabia ($3bn)	HUMAIN (Public Investment Fund) and Blackstone joint venture; national AI champion campus

See adapted E4B Vancouver system guidance

The Gulf Cooperation Council (GCC) has emerged as one of the fastest-growing data centre markets globally, driven by sovereign mandates that have committed over $30 billion between 2025 and 2030. Saudi Arabia established HUMAIN—owned by the Public Investment Fund (PIF)—as a national AI champion and signed an approximately $3 billion partnership with a Blackstone-backed operator for an initial data centre campus. In the UAE, the Stargate UAE project—developed by G42 with OpenAI, Oracle, NVIDIA and SoftBank—targets 5GW of total capacity at a site in Abu Dhabi, with phase one delivering 1GW and of which the first 200MW are expected to go live this year. In Qatar, Syntys has committed $1 billion to support the country's first sovereign AI cloud.

The investment extends to smaller GCC states, including the Bahrain Data Center Park. Acute physical limits in the region constrain these ambitions. A 1 GW data centre in Saudi Arabia is estimated to consume somewhere between 1.5 and 5 million cubic metres of water per year, depending on cooling technology, in a location where the municipal supply relies almost entirely on desalination. Achieving ambitious water-use effectiveness (WUE) targets requires dry or hybrid cooling systems, which consume up to 40% more electricity than water-based alternatives. Immersion cooling has been tested as a regional solution, with a 15-year agreement signed in 2025 for two 10MW facilities in Riyadh and Jeddah using immersion technology and supporting rack densities up to 368 kW—among the highest publicly disclosed anywhere.

However, governments have yet to publish aggregate water demand projections for the data centre sector or relevant sites.

The Global South faces a structural asymmetry. While Africa's data centre market is projected to reach more than $9 billion by 2029, global technology firms dominate the market. In this context, there is an attendant risk of fragmented national regulatory approaches that contribute to “regulatory arbitrage”, with individual governments unable to negotiate fair investment terms from dominant hyperscalers.29
This situation has been referred to as a form of “digital colonialism”, with an estimated 80% of intra-African internet traffic passing through servers in Europe. This has profound implications that may not yet be fully appreciated: data extracted from Southern populations, processed by Northern algorithms, and monetised abroad.

The African Union's Data Policy Framework and Digital Transformation Strategy (2020-2030) provides a blueprint for reform and coordination to remedy the situation, but implementation has lagged.[30]

The strategic question for sovereign wealth funds and development finance institutions is whether GCC-led investment can combine first-mover infrastructure with governance, skills and supply-chain requirements to avoid replicating, in the data centre sector, the dependency patterns that characterise other areas. Climate finance could help bridge the gap, but no dedicated instrument earmarked for data centre resilience in the Global South currently exists.[31,32]

Data centres reject nearly all of the electricity they consume as heat. In northern Europe, operators have begun converting this thermal liability into a revenue stream.

A joint project near Helsinki between a major hyperscaler and a district energy provider is expected to supply approximately 40% of district-heating demand for impacted municipalities, serving approximately 250,000 users. The model has spread across the Nordics. Stockholm Exergi's Open District Heating initiative, launched in 2014, has connected over 25 heat suppliers—including 16 data centres—to its 3,000 km network, recovering enough energy to heat more than 31,000 modern apartments. The European Data Centre Association has reported that 997 data centres in the EU-28 sit within 2 km of a district heating network.[33]

AI-optimised digital twin technology has advanced from concept to deployment. A consortium including NVIDIA unveiled the first digital twin capable of simulating AI data centre power requirements from grid to chip level, using the NVIDIA "Omniverse" platform. The tool enables chip-level power-consumption modelling, what-if scenario analysis, and predictive maintenance scheduling. A cross-sectional survey of 220 professionals, published in the ASRC Conference Proceedings, found that AI-enabled digital twin capability can predict maintenance effectiveness and energy optimisation, with organisational readiness and analytics as essential preconditions.[34]

The Climate Bonds Initiative has documented a 170% annual increase in aligned green asset-backed securities (ABS)—securities collateralised by data centre lease revenues—reaching $7 billion in the first half of 2025 and comprising almost half of all aligned green ABS globally.35
The European Green Bond Standard (EU GBS) establishes a voluntary standard that encourages at least an 85% allocation of proceeds to EU Taxonomy-aligned activities, providing a disclosure framework for operators seeking to demonstrate that green finance commitments are matched by verifiable environmental performance.[36]

The data centre sector and the energy transition are now bound together. Capital commitments have outpaced the grids, water systems and regulatory regimes required to absorb them. The gap between announced capacity and deliverable infrastructure continues to widen. For economies that produce energy and export capital—those of the Gulf and the broader Global South in particular—the key question is no longer participation but rather the questions of who governs the infrastructure, who bears the environmental cost and who retains lasting value.

The risks are significant and correlated. Climate-related costs across the existing stock of data centres may reach $3.3 trillion by 2055 under a high-emissions scenario, mostly driven by heat and drought. Chip replacement cycles of months conflict with debt maturities of years, creating stranded-asset risks. Transmission failures have demonstrated the hazards that fragile grids pose to healthcare, financial and public-service platforms that rely on an uninterrupted cloud supply. Meanwhile, research evidence shows that per-unit efficiency gains can stimulate additional demand in a way that raises aggregate energy consumption.

Public disclosure has not kept pace. Fewer than half of EU data centres have met initial EED reporting obligations and two-thirds of infrastructure asset owners have conducted no independent assessment of physical climate exposure. A credible European rating framework for data centres is unlikely to be in place before mid-2027. Where equivalent standards do not exist, the asymmetry invites "regulatory arbitrage".

Yet there are tangible opportunities. Green securities collateralised against data centre revenues exceeded $7 billion in the first half of 2025. Recovered server heat now supplies district networks at scale in northern Europe. AI-driven digital twins have yielded measurable gains in maintenance and energy performance. There is an emerging architecture for verified data centre commitments, yet adoption remains voluntary. Jurisdictions that pursue grid reinforcement, community consent for data centres, transparent energy and water reporting, and incorporate AI load forecasts in national planning, will secure durable investment. Those that delay enabling policy measures will forgo resources, entrench dependency and forfeit the investment opportunity. Across Africa—where international firms hold above 70% of data centre capacity and four-fifths of intra-continental traffic still transits European infrastructure—improved investment terms for data centres are a necessity.

1.International Energy Agency (IEA). Spencer Thomas and Singh Siddharth. Energy and AI: World Energy Outlook Special Report. April 2025. Available at: https://iea.blob.core.windows.net/assets/b8a83930-5c77-4da7-b795-270ab6a6c272/EnergyandAI.pdf [accessed 24 February 2026].

2.International Gas Union. The role of gas in powering AI-driven energy demand. April 2025. Available at: https://www.igu.org/press-releases/gas-for-data-centres-report [accessed 24 February 2026].

3.Einhorn, G. and King, D. World Economic Forum (with S&P Global data). The $3.3 trillion climate question: can data centres take the heat? October 2025. Available at: https://www.weforum.org/stories/2025/10/data-centres-3-3-trillion-question-heat-cooling/ [accessed 24 February 2026].

4.XDI Systems. 2025 Global data centre physical climate risk and adaptation report. July 2025. Available at: https://xdi.systems/news/2025-global-data-centre-physical-climate-risk-and-adaptation-report [accessed 24 February 2026].

5.McKinsey & Company. The cost of compute: a $7 trillion race to scale data centers. April 2025. Available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers [accessed 24 February 2026].

6.ENTSO-E. 28 April 2025 Blackout. April 2025. Available at: https://www.entsoe.eu/publications/blackout/28-april-2025-iberian-blackout/ [accessed 24 February 2026].

7.Data Center Dynamics, citing NERC incident report. Virginia narrowly avoided power cuts when 60 data centers dropped off the grid at once. 19 March 2025. Available at: https://www.datacenterdynamics.com/en/news/virginia-narrowly-avoided-power-cuts-when-60-data-centers-dropped-off-the-grid-at-once/ [accessed 24 February 2026].

8.Reuters. In Australia, a data centre boom is built on vague water plans. 15 September 2025. Available at: https://www.reuters.com/sustainability/land-use-biodiversity/australia-data-centre-boom-is-built-vague-water-plans-2025-09-15/ [accessed 24 February 2026].

9.Office of the Governor of California. Assembly Bill 93 veto message. 11 October 2025. Available at: https://www.gov.ca.gov/wp-content/uploads/2025/10/AB-93-Veto.pdf [accessed 24 February 2026].

10.Data Center Watch. Data Center Watch report Q2 2025 update. June 2025. Available at: https://www.datacenterwatch.org/q22025 [accessed 24 February 2026].

11.Kyndryl. Fiscal 2025 TCFD report. 2025. https://www.kyndryl.com/content/dam/kyndrylprogram/doc/en/2025/fiscal-2025-tcfd-report.pdf [accessed 24 February 2026].

12.Uptime Institute. 15th Annual Global Data Center Survey. 30 July 2025. Available at: https://uptimeinstitute.com/resources/research-and-reports/uptime-institute-global-data-center-survey-results-2025 and EED commentary https://journal.uptimeinstitute.com/the-energy-efficiency-directive-requirements-come-into-focus/ [accessed 24 February 2026]

13.European Commission. Delegated Regulation (EU) 2024/1364 on data centre sustainability reporting. Entered into force 6 June 2024: https://eur-lex.europa.eu/eli/reg_del/2024/1364/oj/eng [accessed 24 February 2026].

14.European Commission. EU AI Act (Regulation 2024/1689) — guidelines on obligations for general-purpose AI providers. 2025. Available at: https://digital-strategy.ec.europa.eu/en/faqs/guidelines-obligations-general-purpose-ai-providers [accessed 24 February 2026].

15.UK Department for Science, Innovation and Technology (DSIT). Data centres — Cyber Security and Resilience Bill factsheet. November 2025. Available at: https://www.gov.uk/government/publications/cyber-security-and-resilience-network-and-information-systems-bill-factsheets/data-centres [accessed 24 February 2026].

16.World Meteorological Organization (WMO). Advancing early warning systems for resilient energy. 16 December 2025. Available at: https://wmo.int/media/news/advancing-early-warning-systems-resilient-energy [accessed 24 February 2026].

17.Wu, J. (2025) 'Digital Jevons paradox in urban data center energy systems', Nature Cities, 2(8). Available at: https://www.nature.com/articles/s44284-025-00289-9 [accessed 24 February 2026].

18.Mhlanga, D. (2025) 'AI beyond efficiency, navigating the rebound effect in AI-driven sustainable development', Frontiers in Energy Research, 13. Available at: https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2025.1460586/full [accessed 24 February 2026].

19.Vertiv and NVIDIA (2023) 'Power usage effectiveness analysis of a high-density air-liquid hybrid cooled data center', Journal of Electronic Packaging, 145(3). New York: American Society of Mechanical Engineers. Available at: https://www.vertiv.com/en-us/about/news-and-insights/articles/blog-posts/quantifying-data-center-pue-when-introducing-liquid-cooling/ (Accessed: 24 February 2026).

20.Microsoft Sustainability Report 2024, https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/msc/documents/presentations/CSR/Microsoft-2024-Environmental-Sustainability-Report.pdf [accessed 24 February 2026].

21.European Commission Directorate-General for Energy. Minimum performance standards for EU data centres. 2025. Available at: https://energy.ec.europa.eu/resources/preparatory-studies/minimum-performance-standards-eu-data-centres_en and for CADA notification https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/779251/EPRS_BRI(2025)779251_EN.pdf [accessed 24 February 2026].

22.DIGITALEUROPE. Response to EUDCEAR's second technical report and recommendations on revising the reporting scheme. January 2026. Available at: https://cdn.digitaleurope.org/uploads/2026/01/Greentransition_59319_Approved_DIGITALEUROPE_Position_paper_on_rating_and_reporting_scheme_for_data_centres.pdf [accessed 24 February 2026].

23.Commission for Regulation of Utilities (CRU), Ireland. Decision on new electricity connection policy for data centres. December 2025. Available at: https://www.cru.ie/about-us/news/the-cru-publishes-its-decision-on-new-electricity-connection-policy-for-data-centres/ [accessed 24 February 2026].

24.IFRS Foundation. IFRS Foundation publishes guidance on disclosures about transition plans. 31 May 2025. Available at: https://www.ifrs.org/news-and-events/news/2025/06/ifrs-publishes-guidance-disclosures-transition-plans/ [accessed 24 February 2026].

25.EDHEC Infrastructure & Private Assets Research Institute. Physical climate risk survey. 2024. Available at: https://climateinstitute.edhec.edu/news/physical-climate-risk-survey [accessed 24 February 2026].

26.Federal Energy Regulatory Commission (FERC). Order on tariff reform for co-located large loads, PJM Interconnection. 18 December 2025. https://www.ferc.gov/news-events/news/ferc-directs-nations-largest-grid-operator-create-new-rules-embrace-innovation-and [accessed 24 February 2026]

27.Oxford Economics. UK data centre electricity demand forecast. 2025. https://www.oxfordeconomics.com/resource/the-uks-data-centre-boom-growth-trends-drivers-and-the-rising-power-challenge/ [accessed 24 February 2026]

28.Energy Institute. https://knowledge.energyinst.org/new-energy-world/article?id=138649 [accessed 24 February 2026]

29.New America. Africa's digital sovereignty trap: the data center dilemma. 30 July 2025. Available at: https://www.newamerica.org/planetary-politics/briefs/africas-digital-sovereignty-trap/ [accessed 24 February 2026].

30.De Freitas, M.V. (2025) 'Digital sovereignty and data colonialism: shaping a just digital order for the Global South', Policy Center for the New South, PP 38-25, October. Available at: https://www.policycenter.ma/sites/default/files/2025-10/PP_38-25.pdf [accessed 24 February 2026].

31.Stern, N. et al. (2025) 'Delivering an integrated climate finance agenda in support of the Baku-to-Belém roadmap', LSE Grantham Research Institute on Climate Change and the Environment, November. Available at: https://www.lse.ac.uk/granthaminstitute/publication/delivering-an-integrated-climate-finance-agenda-in-support-of-the-baku-to-belem-roadmap/ [accessed 24 February 2026].

32.Middle East Institute. AI, the Gulf and the US: a primer. February 2026. Available at: https://mei.edu/report/ai-the-gulf-and-the-us-a-primer/ [accessed 24 February 2026].

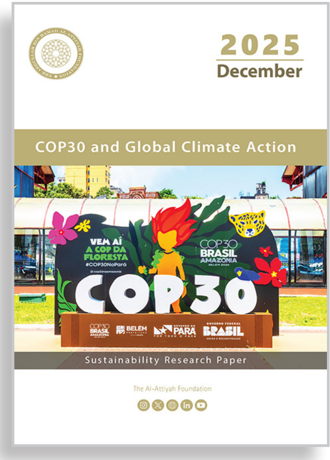
33.European Data Centre Association (EUDCA). An introduction to data centre heat reuse (white paper). August 2024. Available at: https://www.eudca.org/an-introduction-to-data-centre-heat-reuse [accessed 24 February 2026].

34.Adesanwo, O. et al. (2025) 'AI-enabled digital twin framework for predictive maintenance and energy optimization in industrial systems', ASRC Conference Proceedings. Available at: https://global.asrcconference.com/index.php/asrc/article/view/72 [accessed 24 February 2026].

35.Climate Bonds Initiative. Sustainable Debt: Global State of the Market H1 2025. 2025. Available at: https://www.climatebonds.net/files/documents/publications/Sustainable_Debt_H1_2025_01C.pdf [accessed 24 February 2026].

36.ISS Corporate. EU Green Bond Standard sees strong 2025 momentum. 10 November 2025. Available at: https://www.iss-corporate.com/resources/blog/eu-green-bond-standard-sees-strong-2025-momentum/ [accessed 24 February 2026].

Have you missed a previous issue? All past issues of the Al-Attiyah Foundation's Research Series, both Energy and Sustainability Development, can be found on the Foundation's website at www.abhafoundation.org/publications

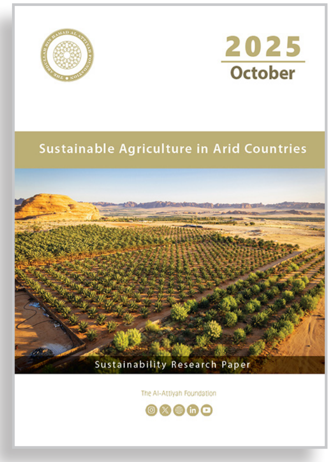


December – 2025

COP30 and global climate action

The 2025 United Nations Climate Change Conference (COP30) in Belém concluded with outcomes that underscored a structural shift in climate governance: non-state actors such as cities, businesses, and civil society brought forward implementation pipelines while national government commitments lagged.

(QRCO.DE)

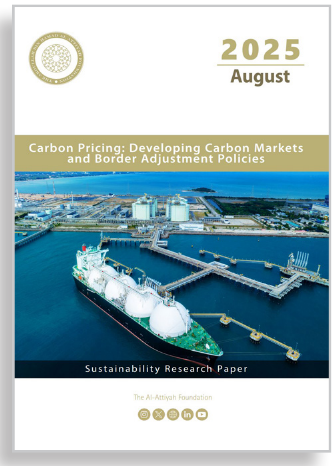


October – 2025

Sustainable agriculture in arid countries

MENA holds 1% of global freshwater yet sustains 6% of the global population, with most of the world's water-stressed countries located in the region.

(QRCO.DE)



August – 2025

Carbon Pricing: Developing Carbon Markets and Border Adjustment Policies

Carbon pricing through taxes, cap-and-trade or crediting, attaches a cost to greenhouse-gas emissions and rewards low-carbon choices.

(QRCO.DE)

Our partners collaborate with The Al-Attiyah Foundation on various projects and research within the themes of energy and sustainable development.

THE ABDULLAH BIN HAMAD AL ATTIYAH FOUNDATION

The Al-Attiyah Foundation

Tel: +(974) 4042 8000,
Fax: +(974) 4042 8099
www.abhafoundation.org

Barzan Tower, 4th Floor, West Bay.
PO Box 1916 Doha, Qatar
Alattiyahfndn

AlAttiyahFndn
The Al-Attiyah Foundation
Al-Attiyah Foundation