From innovation to overshoot:

How data centre expansion risks derailing climate goals

September 2025

About ECOS

ECOS — Environmental Coalition on Standards is an international NGO with a network of members and experts advocating for environmentally friendly technical standards, policies and laws. We ensure the environmental voice is heard when they are developed and drive change by providing expertise to policymakers and industry players, leading to the implementation of strong environmental principles.

About Open Future

Open Future is a European think tank that develops new approaches to an open internet that maximise societal benefits of shared data, knowledge, and culture. The organisation creates strategies for Digital Commons — democratically governed, collectively managed resources that provide an alternative to traditional ownership models. Open Future focuses on reimagining openness to foster a more balanced digital future that serves the public interest.

Authors

Lead author:

Anastasia Tsougka, Programme Manager, Environmental Coalition on Standards (ECOS)

Contributing author:

Zuzanna Warso, Director of Research, Open Future Foundation

Edited by:

Alison Grace, Senior Press & Communications Manager, Environmental Coalition on Standards (ECOS)

Graphic design:

JQ&ROS Visual Communications, jqrosvisual.eu

Contents

Executive summary and recommendations

Digital technologies are a battleground. Digital infrastructure is increasingly seen as a strategic asset, directly linked to nations' ability to assert geopolitical influence and lead in innovation, competitiveness, security, and supply chain resilience. As a result, political support for artificial intelligence (AI) infrastructure is extensive, further fuelling its growth.

Digitalisation is framed as a vehicle for progress, but its unchecked evolution also has a darker side. It risks undermining climate goals, destabilising energy systems, and deepening environmental and social inequalities. Expanding AI infrastructure has become one of the defining environmental challenges of the digital age.

The increasing popularity of AI and other digital technologies in both private and public spheres is driving a boom in the need for data centres. The International Energy Agency's (IEA) Global Energy Review 2025 explicitly identified data centres driven by AI as a major structural factor in 2024's 4.3% global electricity demand growth.[1] Data centres were also a key contributor to increased electricity use in the buildings sector, accounting for nearly 60% of electricity demand growth last year.

Financial support for AI and data centre expansion in Europe has been widespread,[2] but debate around the environmental footprint, energy consumption, and broader societal impacts of this transition has not. The EU and other governments and industry stakeholders have committed to increasing transparency around the consumption of data centres and AI — and much research and innovation exists to help make AI models and equipment more efficient. However, the dangers are often downplayed, deprioritised, or disregarded.

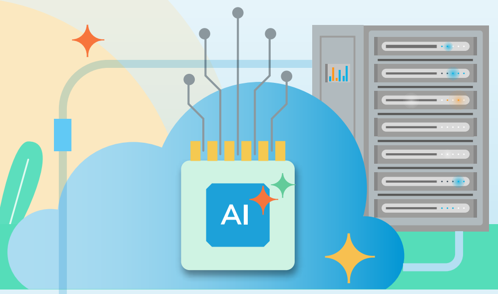
Policymakers are struggling to keep up with the pace of growth. In Europe, developments like the Draghi report brought these discussions to the fore. However, the EU currently lacks a coherent vision or strategy to promote, develop, and regulate AI and its infrastructure — and this comes with risks: environmental, social, and legal.[3]

The lack of balance risks creating a technological landscape in which innovation and growth overshadows sustainability and the public interest, undermining ambitions for a green future. The EU's 2030 climate and renewable energy goals are in jeopardy unless ambitious and enforceable sustainability targets are introduced as part of a coherent EU vision and strategy on the innovation and societal changes driving demand for data centres. Targets without enforceable limits are weak and risk undercutting Europe's progress on climate, energy, and equity.

Until now, efforts to make data centres more sustainable have prioritised efficiency — and while this is needed, it is not the only solution. We urgently need to go beyond efficiency and confront the structural drivers of digital overconsumption. In this report, we demonstrate how. In Part 1, we present the different kinds of data centres out there, and what impact AI has on them. In Part 2, we elaborate on what environmental ambition looks like for this sector, and why we need it. In Part 3, we summarise the existing EU policy landscape — and the opportunities to course-correct — concluding in Part 4 with recommendations for action.

As policymakers play catch-up and new EU initiatives take shape, there is a unique window to align digital developments with environmental and social responsibility. Integrating principles such as sufficiency, circularity, and transparency into the foundations of digital infrastructure will help to ensure that technological progress remains compatible with planetary boundaries and public interest — driving a digital transition that works for us all.

Recommendations



A coherent EU vision and strategy addressing the innovation and societal changes driving demand for data centres is vital. We recommend that the EU develops a holistic approach to the drivers and impact of data centre uses, including ambitious environmental targets, as well as geostrategic, competitiveness, and societal objectives.

1 Focus on more than just efficiency

- **Introduce legally binding, sector-specific targets for reducing energy consumption and greenhouse gas emissions** in data centres, aligned with EU climate goals.
- **Incentivise data centres to source renewable electricity**, applying strict additionality criteria to ensure only newly added renewable capacity is counted.
- **Integrate data centres into national energy strategies** to mitigate grid stress and avoid fossil fuel lock-in.
- **Establish targets for circularity across the full lifecycle of data centre infrastructure**, from design to decommissioning.

2 Prioritise transparency

- **Develop regulatory and standardisation tools to guide the sustainable development** of data centres by establishing mandatory, harmonised methodologies for reporting data centre environmental impacts, including energy, water, emissions, and materials, to increase transparency and prevent greenwashing.
- **Exclude Guarantees of Origin (GOs)/Renewable Energy Certificates (RECs) from environmental reporting** if they are used to offset rather than avoid actual emissions.
- **Make it a condition of public funding and incentives** to demonstrate progress toward environmental and social objectives. Policies must reward actual reductions in environmental impact, not specific technologies or sectors.

3 Adopt sufficiency principles

- **Optimise existing infrastructure** to ensure current data centres operate at full potential before approving new construction. Expansion must be justified, with particular caution taken in regions that face energy constraints.
- **Prioritise sufficiency measures** by optimising digital workloads, minimising non-essential data storage, and focusing AI and digital services on socially valuable applications.

4 Give society a choice

- **Involve local communities and stakeholders** from the outset of any new data centre project. Conduct inclusive consultations and supply chain-wide impact assessments, with special attention to marginalised and vulnerable groups.
- **Develop consumer-facing awareness campaigns to inform users of the environmental impacts of digital services** and promote more sustainable digital habits. Introduce labelling for AI and digital services where relevant.

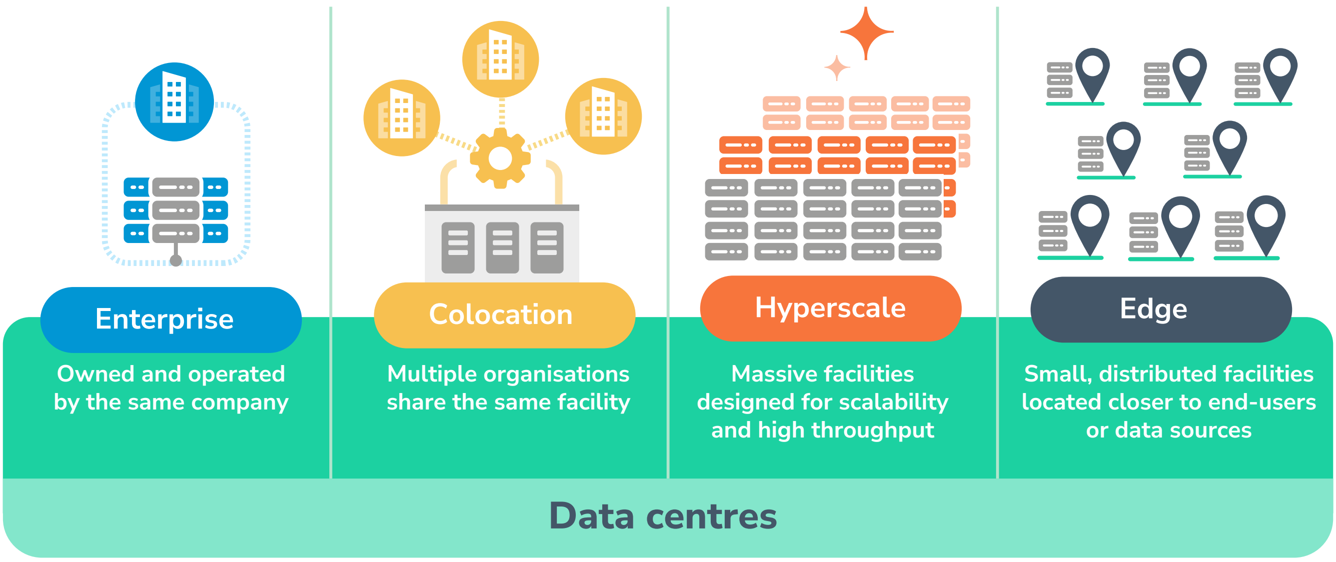
Types of data centres and AI's impact on them

There are many different types of data centres; however, the EU lacks common definitions and indicators for different data centre types that would help it to regulate the sector, as this 2021 ECOS-Coolproducts paper shows.[4] This hinders the EU's ability to properly address the growing climate impacts of AI and data centre expansion. We summarise the main definitions below with more details provided in Annex I.

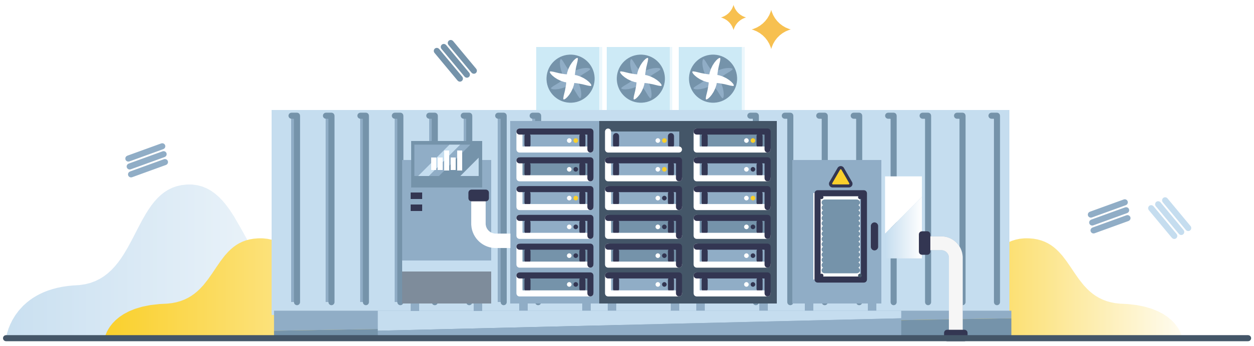
What is a data centre?

A data centre is a specialised facility that houses IT systems and data. It contains large numbers of computer servers, storage devices, and networking equipment. In essence, it is the physical backbone that powers software applications and services by supplying computing power and data storage. Data centres require uninterruptible power supplies, backup generators, and cooling mechanisms to keep hardware at safe operating temperatures.

Modern data centres come in various forms designed for different needs. A single organisation might even use multiple types of data centres to support different workloads. The major categories include **enterprise, colocation, hyperscale, and edge data centres.**



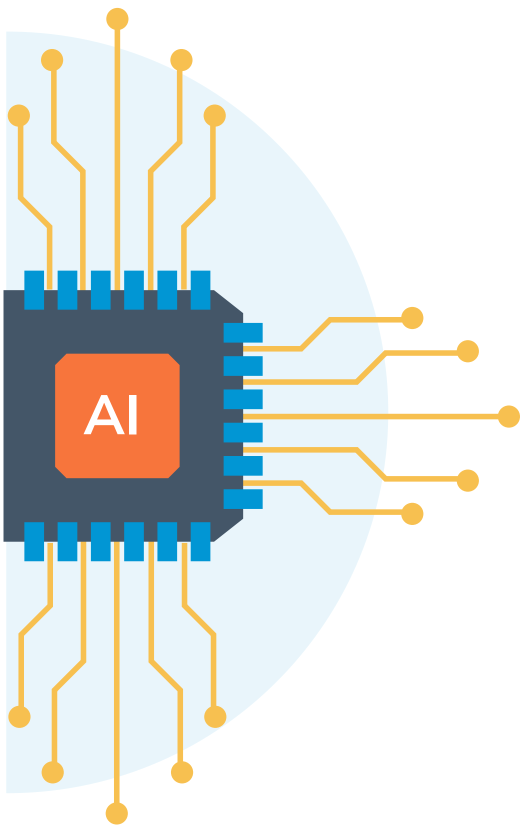
How AI shapes data centre use



The rise of AI is reshaping data centre demand, design, and usage. AI workloads — especially training complex machine learning models — demand enormous computational power and specialised hardware. As a result, certain categories of data centres have evolved or been repurposed specifically to support AI.

The major cloud platforms and tech companies have built AI-optimised clusters within their hyperscale data centres. These are cloud data centres outfitted with thousands of accelerators (like Graphics Processing Units, GPUs, or Tensor Processing Units, TPUs) to train and serve AI models for users worldwide. For example, Microsoft created a dedicated supercomputer for OpenAI hosted in its Azure cloud, featuring over 285,000 central processing unit (CPU) cores and 10,000 GPUs all linked by an ultra-fast 400 Gb/s network. In hyperscale AI centres, the infrastructure is scaled out massively: companies interconnect racks upon racks of GPU servers with high-bandwidth fabrics so they can work in parallel on AI training.

High-performance computing centres (HPCs) are also adapted to fit AI needs. HPCs typically house supercomputers — large clusters of nodes with powerful CPUs, specialised accelerators, and high-speed interconnects — used for scientific simulations, modelling, and increasingly, machine learning. Many HPC centres have added GPU partitions or AI accelerators to accommodate researchers training neural networks. For example, the Ohio Supercomputer Center recently deployed a new GPU cluster ("Ascend") with 96 Nvidia A100 GPUs connected by Nvidia's 200 Gb/s InfiniBand network, delivering an extra 2 petaflops of AI computing capacity to scientists.[5] Such HPC systems can run AI workloads at scale similarly to industry hyperscale clusters. The key difference is that HPC centres are typically publicly funded or academic, and they cater to advanced research in fields like medicine, climate modelling, or physics in addition to AI. With AI's growing importance in research, the line between traditional HPC and AI supercomputing is blurring.



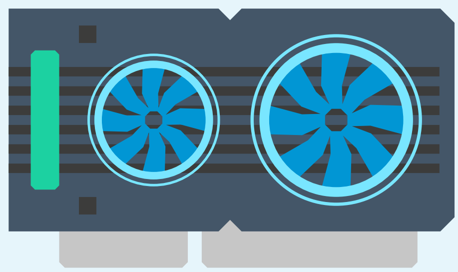
Finally, purpose-built AI complexes, also known as **AI supercomputing facilities**, are created by both **companies** and governments more frequently. For instance, Meta built the AI Research SuperCluster (RSC),[6] a private AI supercomputer housed in a data centre, with 6,080 Nvidia A100 GPUs initially and scaling to 16,000 GPUs. Running AI algorithms — especially deep learning — efficiently at scale requires specialised data centre hardware and architecture. In contrast to standard enterprise applications, AI workloads involve parallel processing of huge datasets and model computations. Data centres used for AI therefore incorporate AI-focused infrastructure components that significantly boost performance. Those are mainly GPU and TPU clusters, AI accelerators, high-speed networking, and liquid cooling solutions.

Graphics Processing Units (GPUs) have become the workhorses for AI training due to their ability to perform many computations in parallel. AI data centres deploy servers with multiple high-end GPUs connected together. These GPU clusters can be scaled to hundreds of thousands of chips working in unison on a single training job. Google's data centres alternatively use **Tensor Processing Units (TPUs)**, which are designed specifically for machine learning tasks. Beyond general-purpose GPUs, companies are now using specialised **AI accelerators**. These are chips architected from the ground up for machine learning workloads. To make thousands of AI chips work together effectively, the connections between servers are just as critical as the chips themselves. AI data centres use high-speed network fabrics with minimal latency.

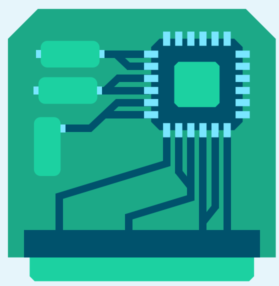
The dense packing of AI hardware and the heavy, sustained processing load mean these data centres generate a lot of heat. Traditional air cooling (blowing cold air through racks) is often insufficient or inefficient for high-performance AI clusters. As a result, many AI data centres are adopting liquid cooling as an efficiency solution, which entails circulating chilled water or coolant directly to hot components (CPUs, GPUs) via cold plates, or even immersing entire servers in dielectric fluid. This innovative technology can help reduce some of the energy consumption associated with energy-intensive air-cooling operations, as well as direct water consumption.

The solutions required for the optimal functioning of high-tech data centres are resource-intensive and require massive investments in infrastructure.

Graphics Processing Unit (GPU)



Tensor Processing Unit (TPU)



Environmental and social impacts of AI-driven data centres

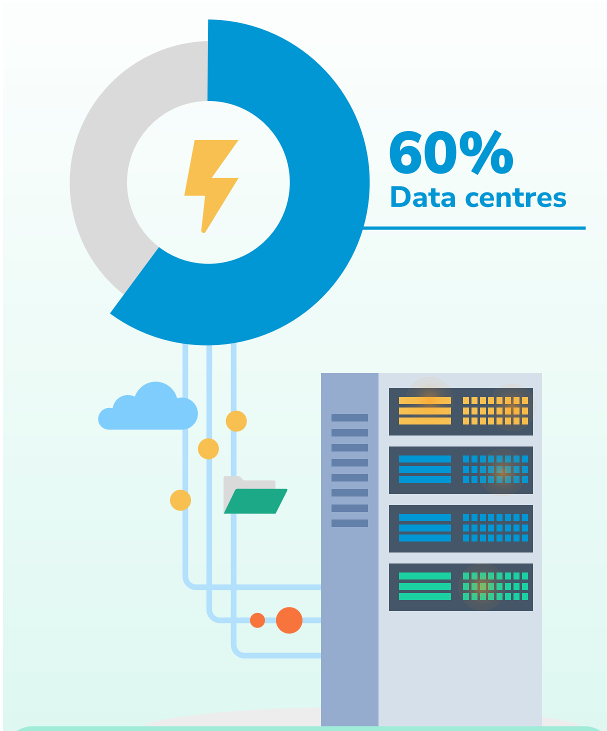
The exponential growth of AI and data centres raises significant environmental and social concerns that extend well beyond energy consumption. We summarise the main impacts in this section.

Rising environmental concerns

Data centres are essential infrastructure for the digital economy. However, their rapid expansion is driven by an increasing demand for computational power — often dedicated to tasks of questionable necessity or low societal value. AI, although not the sole contributor, exacerbates this issue by fostering a culture of computational excess.

Large-scale AI models require immense processing capabilities, often leading to inefficient resource use and additional strain on power grids. Both the training and inference phases of AI models demand enormous computational power, contributing to high carbon emissions and resource consumption — including resource extraction for hardware production and e-waste generation.

The EU needs a **strategic and holistic approach to measuring and mitigating environmental impacts across the entire data centre and AI lifecycle** — covering both software and hardware stages.



Data centres were responsible for 60% of the growth in electricity demand in the buildings sector in 2024

Source: IEA

The difference between training and inference

What are training algorithms?

Training algorithms are processes that enable an AI model to learn from large datasets. By repeatedly adjusting internal parameters, the model identifies patterns and relationships in the data to perform specific tasks, such as recognising images or understanding language. This phase is highly resource-intensive, often requiring specialised hardware (e.g. Graphics Processing Units, GPUs, or Tensor Processing Units, TPUs), massive computing power, and significant energy use.

What are inference algorithms?

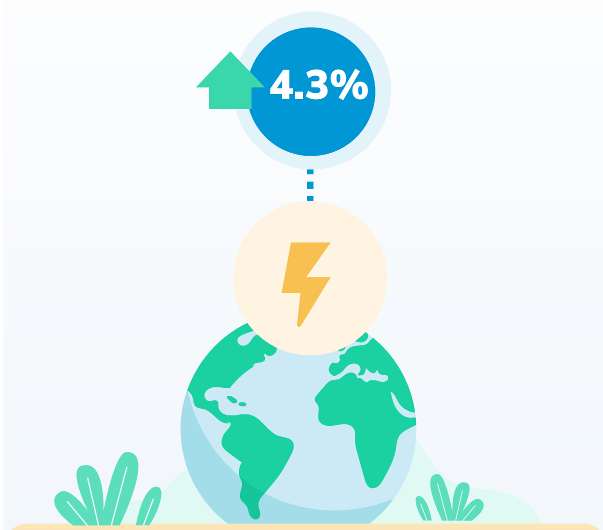
Inference algorithms are used after a model has been trained. They apply the model's learned knowledge to new data to make predictions or decisions — such as classifying a photo or answering a user query. In general, inference consumes less energy and computing power than training. However, the resource demand of inference can vary considerably depending on the model's size, the type of input (e.g., long or complex prompts), and the task's requirements (e.g., real-time processing or multimodal output). At large scale, inference can still be a major contributor to overall energy consumption.

Increasing energy demand and its risks

Energy demand is rising

Data centres are placing growing pressure on electricity systems and complicating efforts to decarbonise energy, making them a critical challenge in the broader energy transition.

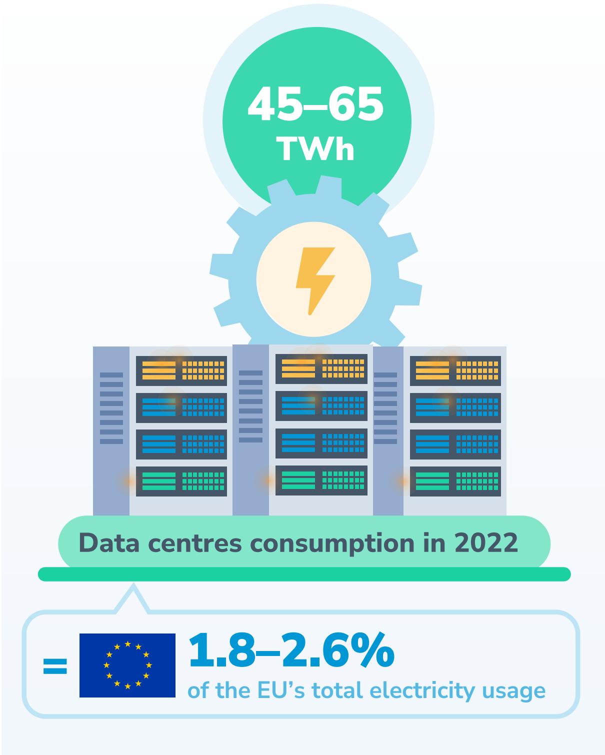
The International Energy Agency's (IEA) Global Energy Review 2025 explicitly identified data centres driven by AI as a major structural factor in 2024's 4.3% global electricity demand growth.[7] Data centres were also a key contributor to increased electricity use in the buildings sector, accounting for nearly 60% of electricity demand growth last year. The IEA also reports that the installed capacity of data centres grew in 2024 by ~20% (≈15 GW), mostly in the United States and China. These data centres are still largely powered by fossil fuels, which still generate nearly 60% of electricity globally — and coal alone providing 35%.



The 4.3% growth in global energy demand in 2024 was due largely to AI

Source: IEA

According to older data from the IEA, global data centre electricity consumption in 2022 was estimated to be between 240 and 340 terawatt-hours (TWh), accounting for approximately 1-1.3% of global final electricity demand.[8] This figure excludes cryptocurrency mining, which adds roughly another 110 TWh, or 0.4% of global demand. Some analyses accounting for emerging AI workloads suggest the total could be on the higher end; for example, one IEA analysis put 2022 data centre usage around 460 TWh (nearly 2% of global electricity) when including recent growth in AI and other power-intensive computing.[9]



A 2024 European Commission report estimated that EU data centres consumed about 45–65 TWh in 2022, which is roughly 1.8–2.6% of the EU's total electricity usage.[10] Some smaller countries with major cloud/hyperscale hubs have an outsized share — notably Ireland, where data centre power demand has more than tripled since 2015 and now represents approximately 18% of its total electricity use. Other countries with high data centre shares include the Netherlands (~5% of national electricity), Luxembourg (~4.8%), Denmark (~4.5%), and Sweden (~2.3%) — compared to ~2–3% in larger EU economies like Germany and France.

A 2024 report by McKinsey projects that Europe's data centre power consumption is expected to triple by 2030, rising from approximately 62 terawatt-hours (TWh) to over 150 TWh. This surge would elevate data centres' share of total European electricity consumption from 2% to around 5%.[11]

Individual hyperscale data centres can draw enormous power.[12] These facilities typically have power capacities of tens of megawatts (MW), with the largest exceeding 100 MW of electrical draw. For perspective, a data centre with a continuous 100 MW power usage (not uncommon for a big cloud campus or colocation hub) would consume roughly 876,000 MWh (876 GWh) of energy in a year, which is about 0.8 TWh annually. This is comparable to the yearly electricity use of approximately 80,000 U.S. households.

Artificial intelligence applications, especially large-scale models, contribute significantly to this energy demand. AI has become increasingly popular for low-effort tasks and is advertised as the answer to improved productivity in the workplace. Its widespread deployment tends to prioritise commercial and non-essential applications — for example, making videos, photos, or deepfakes for social media. The IEA has noted that the rapid and mainstream adoption of AI technologies, such as chatbots, is likely to further accelerate energy consumption growth.

AI is not the only culprit

While significant, AI is not the only culprit for the increasing energy consumption of data centres. Data centres are working overtime to perform multiple energy-intensive tasks and provide other energy-intensive digital services.

For instance, Bitcoin mining alone consumes an estimated 155 to 172 terawatt-hours (TWh) annually comparable to the energy consumption of Poland.[13] Additionally, streaming platforms such as Netflix and YouTube collectively consume over 200 TWh annually, a substantial portion of which is processed through data centres.[14] Cloud gaming further amplifies energy demands; for example, PC cloud gaming can draw approximately 340 watts per user in data centres, with an additional 180 watts for network transmission.[15] Social media platforms, with over 5 billion users globally, also generate continuous data traffic necessitating extensive processing and storage infrastructure.

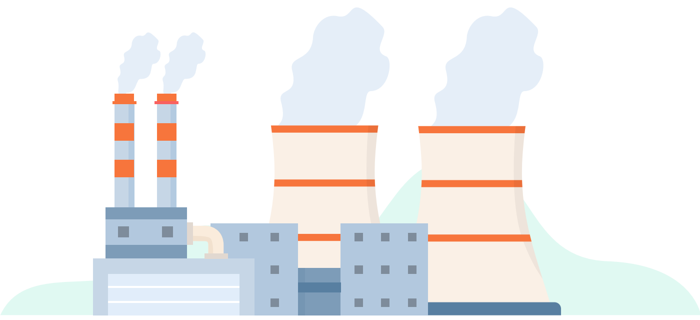
Climate goals are at risk

There is an industry-wide pattern of climate backtracking, as soaring energy use clashes with previously stated environmental ambitions. Several major tech companies have publicly acknowledged their inability to meet their climate goals.

For example, Google reported a 48% increase in emissions since 2019, citing the electricity demands of AI as a major barrier to its 2030 carbon-free energy target.[16] Microsoft saw its emissions rise by nearly 30% since 2020 and has since revised its sustainability strategy to account for the energy-intensive growth of its AI services.[17] Meta's emissions more than doubled since 2019, reaching 14 million metric tonnes of CO_2 equivalent in 2023, largely due to heavy investment in AI infrastructure.[18] Amazon, while reporting a 3% decline in its carbon footprint, was removed from the Science Based Targets initiative for its lack of a credible emissions reduction plan.[19]

The role of nuclear and other sources of power generation

In response to the escalating energy demands of data centres, there has been a notable shift towards exploring new solutions beyond traditional renewables in recent years, despite Big Tech having previously invested in wind and solar.



Big Tech companies have been investing in small modular reactors (SMRs), which would — if ever shown to be feasible — produce nuclear power, to secure low-emissions energy for their data centres. Google recently entered into an agreement with nuclear power start-up Kairos Power to provide low-carbon energy for its data centres and AI operations, aiming to deliver up to 500 megawatts of power from SMRs by 2035.[20] Amazon also announced investments in SMRs to meet the increasing power demands.[21]

The increase in nuclear power generation is confirmed by the IEA's 2025 Global Energy Review. In 2024, nuclear electricity generation rose by 100 TWh, marking the largest annual increase this century outside of the post-COVID rebound. This surge was driven by the addition of over 7 GW of new nuclear capacity, a 33% increase compared to 2023. As of February 2025, there were 62 nuclear reactors under construction across 15 countries, amounting to nearly 70 GW of capacity.

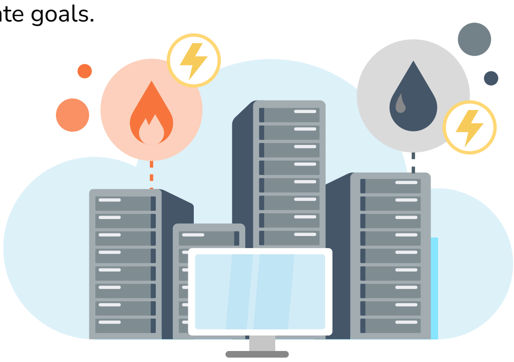
The IEA states that nuclear energy is set to play a crucial role in meeting the increasing electricity demand driven by electrification and digitalisation, predicting that nuclear power generation will reach new heights in 2025 and will continue to rise for the following two years.

These developments indicate a worrying shift towards non-renewable electricity sources to accommodate the substantial energy requirements of modern data processing. While nuclear power is being proposed as a solution to the growing energy demands of data centres and concerns about carbon emissions, few openly discuss the severity of the potential consequences. Nuclear energy comes with important environmental challenges, such as the risk of radioactive waste and contamination as well as wider geopolitical and security risks. Furthermore, newly built reactors often fail to deliver on time and on budget.[22] Finally, the use of nuclear power can overshadow other critical environmental issues, including the significant water consumption and thermal pollution associated with the cooling processes in both the nuclear power plants and the data centres themselves.

Powering data centres: Renewables vs. fossil fuels

Major oil and gas companies have also shown interest in the data centres energy market. ExxonMobil announced plans to construct a 1.5-gigawatt natural gas-fired power plant dedicated exclusively to supplying electricity to data centres.[23] This facility claims it is designed to incorporate advanced carbon capture and storage (CCS) technology, aiming to capture and store over 90% of its carbon dioxide emissions. Chevron similarly announced it would explore suppling low-carbon power to data centres.[24]

Unfortunately, the use of natural gas with carbon capture to respond to the escalating energy demand is another concerning approach. Although it presents itself as an easy, short-term answer to some emissions reduction, it is costly and untested, risking exacerbating the already big reliance of the data centre industry on fossil fuel-based solutions without increasing emissions enough to achieve climate goals.



Data centres already run on grids that are largely fossil fuel-dependent and use fossil fuel-based backup systems, such as diesel generators, to address grid stability issues during peak demand and ensure uninterrupted service. Choosing to continue using natural gas with CCS will lock the industry into long-term carbon dependency, diverting attention and resources from renewable solutions such as wind and solar power. Instead of prioritising genuine sustainability, these efforts will likely reinforce business-as-usual models that perpetuate environmental harm under the guise of low-carbon innovation.

Nevertheless, it is important to underscore that the rapid growth of data centres and their soaring energy demand complicates the energy transition in ways that go beyond phasing out fossil fuels. Even when powered by renewable sources such as solar, wind, or hydro, data centres place significant strain on existing clean energy infrastructure. If technology companies are not investing in expanding renewable supply, they compete for a limited resource — diverting electricity that could otherwise support the decarbonisation of residential, industrial, or transport sectors.

In addition to supply constraints, data centres also exert mounting pressure on electricity grids, which are often not designed to handle such concentrated and continuous demand. This can overwhelm local infrastructure, delay connections for other sectors, and prompt costly upgrades — raising critical questions about who should bear these costs. For example, in the region of Flanders, Belgium, recent reports confirmed that the regional power grid is no longer strong enough to connect all new business users, illustrating how grid capacity, not just energy supply, is becoming a bottleneck in the transition.[25]

The unchecked expansion of AI-driven data centres poses a significant risk to achieving climate targets and progressing in the energy transition, threatening to undermine decarbonisation efforts, if additionality of renewable electricity and grid development are outpaced by the need for more infrastructure and more energy.

Underreporting of emissions

Reports indicate that tech companies significantly underreport their data centre emissions, with actual emissions estimated to be an enormous 662% higher than what is publicly disclosed.[26] The reliance on renewable energy certificates (RECs) allows companies to claim sustainability without addressing the true environmental costs of their operations, with companies often purchasing RECs to claim carbon neutrality, even though their data centres continue to draw power from fossil-fuel-based grids. The opacity surrounding emissions data and the complexity of tracking third-party data centre emissions further complicate efforts to hold tech giants accountable. A recent investigation revealed that many tech companies are using outdated carbon accounting methods to underreport their emissions, obscuring the actual environmental costs of AI expansion.[27]

Offsetting is also used to claim lower emissions, while doing little for climate goals.[28] Microsoft, for instance, recently made a USD 200 million deal with Re.green, a Brazilian start-up, to purchase 3.5 million carbon credits over 25 years.[29] This initiative aims to offset rising emissions from its AI-driven data centre operations by reforesting parts of the Amazon and Atlantic forests. With Microsoft's carbon footprint exceeding 17 million tonnes in 2023, relying on offsets fails to address carbon emissions and real climate impacts. Serious questions have also been raised about many offsetting projects, the majority of which are "likely junk," according to a recent investigation by Corporate Accountability and the Guardian.[30]

Aggravating the water crisis

In addition to energy consumption and carbon footprint, **data centres significantly impact water resources due to their need for substantial amounts of clean water** — used mainly for cooling purposes. Water is consumed throughout the lifecycle of data centre equipment, from the mining of raw materials to semiconductor manufacturing.

On average, a Google data centre consumes approximately 450,000 gallons (approx. 1,703,435 litres) of water per day.[31] A single data centre uses approximately 26 million litres of water annually per megawatt of data centre power. Assuming 20% of all existing European data centres employ liquid cooling systems, the total water usage by data centres would reach 43.2 billion litres annually — enough to fill roughly 173 million Olympic-size swimming pools.[32] This level of consumption can strain local water supplies, particularly in regions already facing water scarcity. A 2024 report highlighted that data centres in Virginia's "data centre alley" (USA) used 1.85 billion gallons of water (approx..7 billion litres), up from 1.13 billion gallons in 2019 (approx. 4.2 billion litres), underscoring the escalating water demands of these facilities.[33]

Advanced AI models also increase heat generation because of the large need for computational power. This also leads to higher water usage for cooling. AI models such as ChatGPT consume about two litres of water for every 10 to 50 queries.[34]

E-waste and raw materials

Beyond water usage, **data centres contribute significantly to electronic waste (e-waste) due to the rapid obsolescence of hardware components**. Equipment such as servers, racks, and circuits are frequently replaced to meet evolving technological demands, leading to substantial e-waste generation. For example, GPUs last approximately 3-5 years. The burden is significant if one considers that in 2019, the world generated 53.6 million metric tons of e-waste — an average of 7.3 kg per capita.[35] In 2024, only 22% of this was properly collected and recycled, according to UNEP, the UN Environment Programme. Increasing circularity and the longevity of ICT products is essential to reduce the need for new materials.

Modern data centres are massive consumers of diverse raw materials, reflecting the scale of physical infrastructure behind our digital world. Data centres depend on a complex flow of metals and elements. High-density server hardware and power systems contain significant amounts of copper (wiring, busbars), aluminium (frames, heat sinks), rare earth magnets (in hard drives and cooling fans), as well as precious and semiconductor materials in circuit boards and chips. This digital infrastructure's material appetite is surging alongside global demand for data.

By 2030, industry analysts predict sharp increases in consumption: for example, AI and data centre growth could add ~1 million tonnes of additional copper demand by 2030, while the shift to lithium-ion UPS is creating a new market for thousands of tonnes of lithium and cobalt in backup systems.[36]

Ultimately, the concept of a digital economy should not be disconnected from that of a circular and green economy. The so-called twin transition in the EU — encompassing both digital and green transformation — must become a real commitment rather than the theoretical concept it is now. Without actionable policies and enforcement, the data centre industry's pursuit of growth will continue to overshadow climate goals, reinforcing existing challenges and putting pressure on planetary boundaries.

The intersection of environmental and social impacts

The rapid expansion of data centres and AI technologies has profound implications not only for the environment but also for communities hosting these data centres and those impacted by activities throughout the entire supply chain.

Because data centres require large amounts of land, electricity, and water to operate, they often strain local resources and exacerbate existing inequalities. The reliance on natural resources and its consequences has a detrimental impact on people's social and economic rights. The water-intensive cooling systems data centres can deplete local water supplies, especially in drought-prone regions. When water resources are diverted to support digital infrastructure, communities can face shortages that affect their access to clean drinking water, agriculture, and sanitation — fundamental human rights recognised by international frameworks. This comes at a time when global demand for freshwater is expected to exceed supply by 40 percent by the end of the decade.[37]

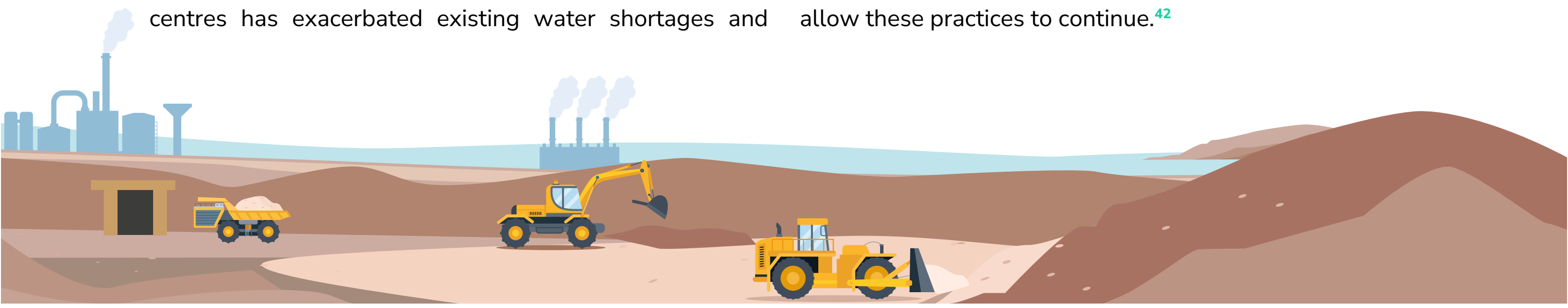
The commodification of natural resources also threatens local small-scale activities, such as traditional farming practices that rely on consistent water access. Often the economic opportunities promised by industrial expansion have not materialised for these communities, leaving them in precarious situations.

Energy consumption also plays a critical role in social and economic inequalities. Increased demand for electricity can drive up energy costs, making it harder for low-income households to afford basic utilities. In regions such as Northern Virginia (USA) and Dublin (Ireland), community groups have protested the establishment of new data centres due to concerns over excessive water use and the strains on already stretched energy grids. In Northern Virginia, residents have reported rising electricity costs and infrastructure strain, as data centres consume nearly as much energy as the entire state of New Hampshire.[38] In Mexico's Querétaro state, the rapid expansion of data centres has exacerbated existing water shortages and placed immense pressure on the electrical grid. Despite claims of sustainability, local communities have voiced frustration over industrial water needs being prioritised over residential supply, highlighting the social inequalities perpetuated by unchecked technological expansion.[39]

In addition, the construction and operation of data centres can contribute to environmental degradation, such as deforestation and land-use change, disproportionately impacting marginalised communities that often depend on natural resources for their livelihoods. Economic disparities are widened as access to AI-powered solutions remains limited to those with the financial means, leaving certain communities undeservingly behind in the digital divide.[40]

Data centres and AI technologies are deeply embedded in global supply chains that extract, manufacture, and distribute components, including minerals, electronics, and data infrastructure. For example, key semiconductor companies, such as NVIDIA, do not manufacture their own GPUs; instead, they rely on a supply chain of mining and refining operations that extract critical minerals from countries such as Brazil, the Democratic Republic of Congo, and Kazakhstan. These mining activities often occur in regions with lax environmental and labour regulations, resulting in extensive ecological degradation and social upheaval. The extraction of essential materials like silicon, copper, and tantalum fuels the growth of AI infrastructure, but it also perpetuates harmful practices, including deforestation, pollution, and the displacement of local communities, highlighting the urgent need for more sustainable and ethical sourcing strategies within the AI industry.[41]

These supply chains encompass a diverse range of workers, from miners to data centre engineers, each facing different working conditions. For example, in Africa and Latin America, the extraction of minerals essential to AI, such as lithium, often involves exploitative labour practices. In Nigeria's Nasarawa region, for instance, illegal lithium mining has skyrocketed, with children working long hours in dangerous conditions for minimal pay. Despite laws prohibiting child labour, weak enforcement and corruption allow these practices to continue.[42]



EU policy overview
a fragmented approach to a growing issue

The European Union currently lacks a coherent strategy covering the entire AI-driven data centre value chain. There are, however, various policies and laws that cover data centres to a greater or lesser extent. There are also some voluntary initiatives that play a role. We review the regulatory landscape below — with its challenges and opportunities.

How data centres are regulated in the EU

There are many overarching laws and voluntary initiatives that could be applied in the context of data centres, but no one piece of legislation sets binding, sector-specific rules to directly contain their environmental impacts and ensure the EU can still reach its climate goals.

The European Commission is charting a path towards climate neutrality by 2050, with a binding target to cut net greenhouse gas emissions by at least 55% by 2030 compared to 1990 levels, as set out in the European Climate Law. To achieve this, the EU calls for existing laws to be fully implemented and industry to be decarbonised through reliance on Europe's existing strengths, such as wind power, hydropower, and electrolysers.

Policy

Energy Efficiency Directive (EED)

The Delegated Act on the First Phase of the EU Data Centre Sustainability Rating Scheme — as requested under Article 12 of the revised Energy Efficiency Directive (EED) (Directive (EU) 2023/1791) — introduces mandatory sustainability reporting for data centres. Data centres with an IT power demand >500 kW must report data on energy consumption, power utilisation, temperature set points, waste heat reuse, water usage, and renewable energy consumption. With collected data published in a publicly accessible, EU-wide database, enabling comparisons and benchmarking. Reporting obligations will gradually expand, with colocation and co-hosting operators required to gather and publish data by 2026.

This Delegated Act can be seen as a first step towards regulating the environmental impacts of data centres. Monitoring and transparent reporting are essential to inform decision-making and increase accountability as well as encourage change. Nevertheless, the EED is a directive, not a regulation, so it leaves a high degree of discretion to EU Member States in setting general efficiency goals and it does not contain an obligation to address the digital sector.

More specifically, under the EED, the EU must reduce final energy consumption by 11.7% by 2030 compared to 2020 projections. Member States must contribute to this target collectively, but there is no obligation to set sector-specific targets — only national targets apply. Because the EED relies heavily on Member States' implementation choices, this creates risks of uneven enforcement. Furthermore, the Directive does not mandate emissions reductions for data centres, leaving the sector's carbon footprint largely unregulated.

During the stakeholder consultation on the delegated act, ECOS supported the overall approach, particularly the use of transparent and multi-layered indicators. However, we strongly criticised the threshold-based scope of the scheme. The final 500 kW threshold excludes a large share of smaller data centres — many of which are less efficient — thereby undermining the representativeness of the reported data and risking loopholes.

Earlier drafts even considered a 100-kW threshold, which ECOS already viewed as too high. We advocated instead for a tiered approach, requiring basic reporting (e.g. Power Usage Effectiveness, PUE, and annual energy use) for all facilities and more detailed data from larger ones. Excluding small operators now sets a precedent that may complicate future efforts to apply minimum performance or labelling requirements to the sector.

Renewable Energy Directive (RED II and RED III)

The Renewable Energy Directive (RED II and its amendment RED III) could partially mitigate these gaps, as it aims to increase the share of renewable energy across all sectors in the EU. Under RED III, the EU must ensure that at least 42.5% of total energy consumption comes from renewable sources by 2030. However, there are no binding or other renewable energy targets specifically for data centres as the sector is not directly addressed in the Directive. Each Member State is required to submit a National Energy and Climate Plan (NECP), outlining how they will contribute to the EU-wide target. As with the EED, this reliance on national discretion weakens the potential for harmonised action across the data centre sector.

Another critical issue, relevant to data centres in the context of RED, is the reliance on Guarantees of Origin (GOs) and Power Purchase Agreements (PPAs) to demonstrate renewable energy use. These mechanisms, while designed to support renewables, are open to manipulation. For instance, companies can purchase GOs without consuming any physical renewable electricity, allowing them to claim "100% renewable" operations despite sourcing electricity from fossil-fuel grids. This creates a paper-based green label rather than real additional renewable energy generation.

In the case of PPAs, companies might sign agreements to support renewable generation in different countries, with no direct link to their actual consumption patterns, yet still report "green electricity use". Both mechanisms allow companies to make green claims, without necessarily reducing fossil fuel dependence or driving real energy transition efforts.

The equivalent in non-EU markets (e.g., the US) are Renewable Energy Certificates (RECs), which operate similarly. Big Tech companies can potentially use RECs to greenwash the energy footprint of their data centres, reinforcing the disconnect between reported and actual emissions reductions.

Ecodesign regulation for servers and data storage products

The Ecodesign Regulation ((EU) 2019/424), currently under review, sets minimum environmental and energy efficiency requirements for servers and data storage products placed on the EU market. This has indirect impacts on data centres, as they will be equipped with servers and data storage products that are more energy and resource efficient. However, the regulation itself does not explicitly impose direct environmental obligations on data centre operators.

Ecodesign requirements cover energy efficiency, repairability, and recyclability, requiring manufacturers to design products so that critical components (e.g., data storage devices, memory, central processing units - CPUs, motherboards, power supply units - PSUs, and chassis) can be replaced. Products must also allow secure data deletion before reuse or recycling. Manufacturers must disclose information on the presence of critical raw materials (e.g., cobalt in batteries, neodymium in hard drives) to assist recyclers. The regulation includes indicative benchmarks for energy efficiency, encouraging best practices.

Overall, this regulation improves the efficiency and circularity prospects of ICT equipment, indirectly benefiting data centre circularity. Nevertheless, meeting minimum standards does not guarantee optimal performance. A data centre operator could still select ICT equipment that only just complies with baseline criteria, bypassing options with far greater energy efficiency or circularity benefits.

CSRD, CSDDD and the EU Taxonomy

The EU's sustainability reporting framework — anchored in the Corporate Sustainability Reporting Directive (CSRD), the Corporate Sustainability Due Diligence Directive (CSDDD), and the EU Taxonomy Regulation — broadly applies to data centre owners, operators, and service providers, particularly in relation to disclosure and transparency obligations.

The EU Taxonomy Regulation, in force since 2020, creates a classification system that defines which economic activities can be considered environmentally sustainable. Companies in scope of the CSRD are required to disclose their alignment with the taxonomy, including for activities such as data processing, storage, and server hosting.

Most large data centre operators, colocation tenants, and cloud service providers are also covered under the CSRD, which mandates the disclosure of Scope 1, 2, and 3 greenhouse gas emissions, energy consumption, and climate-related financial risks. However, the Omnibus Simplification Package, introduced in early 2025, proposes some key changes. It would significantly raise the employee threshold for CSRD applicability — from 250 to 1,000 employees — potentially excluding many data centre operators from mandatory sustainability reporting. In addition, the timeline for compliance is proposed to be extended by two years for companies in the second and third application waves, delaying the point at which climate-related risks and energy performance must be reported.

Under the CSDDD, companies will be required to identify, prevent, and mitigate adverse environmental and human rights impacts across their value chains. While this directive was initially expected to drive upstream accountability — including in ICT supply chains and infrastructure — the Omnibus Package proposes to delay its application until 2028 and to limit due diligence obligations primarily to direct (Tier 1) suppliers. This weakens its potential relevance for operators aiming to address risks linked to hardware manufacturing, critical raw materials, and outsourced digital services.

GPP Criteria for Data Centres, Server Rooms and Cloud Services

The EU Green Public Procurement (GPP) Criteria for Data Centres, Server Rooms and Cloud Services offers a voluntary framework designed to guide public authorities in procuring digital infrastructure with lower environmental impact.[43] The criteria, published in 2020, apply across a broad range of services — including enterprise data centres, small-scale server rooms, cloud platforms (IaaS, PaaS, SaaS), and colocation facilities — and address multiple environmental concerns such as high energy consumption, GHG emissions, hazardous e-waste, use of high-GWP refrigerants, and missed opportunities for waste heat reuse. However, despite its technical robustness, GPP criteria are not legally binding and rely entirely on the willingness and capacity of contracting authorities to implement it. This undermines the potential for systematic, EU-wide progress in reducing the environmental footprint of digital infrastructure.

Structured across four categories — selection criteria, technical specifications, award criteria, and contract performance clauses — the GPP framework allows contracting bodies to choose between a "core" (basic) and "comprehensive" (ambitious) level of ambition. It includes notable requirements such as active state efficiency thresholds for servers), ICT operation range specifications for free cooling, reparability documentation, and energy and cooling monitoring. It also encourages tracking of the Renewable Energy Factor (REF), aiming for full reliance on renewable energy, and limits the global warming potential of refrigerants. Yet these features, while progressive, face shortcomings.

First, their uptake is highly uneven, as many public buyers lack the technical expertise or procurement bandwidth to operationalise criteria that require knowledge of performance standards, energy metrics, and lifecycle assessments. Second, like the RED, the GPP's reliance on proxy tools — such as REF or waste heat infrastructure readiness — can reward declared intentions over verifiable outcomes. There is no enforcement mechanism to ensure that data centres claiming alignment with GPP principles are delivering measurable reductions in energy consumption or emissions.

In short, the GPP criteria for data centres offer a technically sound but politically soft instrument. Without legal obligations or coordinated monitoring, their impact is constrained to isolated cases of best practice rather than driving systemic change in the environmental performance of digital infrastructure.

The Cloud and AI Development Act and the AI Continent Plan

The Cloud and AI Development Act, expected in late 2025 or early 2026, is a major legislative proposal aimed at rapidly scaling up Europe's cloud and AI infrastructure. According to the European Commission's Call for Evidence, the initiative is designed to address strategic gaps in computational capacity, reduce Europe's dependence on foreign cloud providers, and ensure that infrastructure growth is geographically balanced, secure, and sustainable. It proposes to triple EU data centre capacity within 5–7 years, improve access to land, water, and energy, and streamline permitting processes that are currently seen as fragmented and slow.

This initiative explicitly links to other policy pillars, including the AI Continent Action Plan and upcoming Strategic Roadmap for Digitalisation and AI in the Energy Sector (as part of the Affordable Energy Plan), forming part of a broader vision for digital sovereignty and green industrial transformation.

ECOS is concerned that the Commission's framing remains overly focused on enabling growth and investment, with no binding sustainability safeguards currently foreseen. While the Call for Evidence acknowledges challenges such as high capital costs, supply constraints, and increased energy and water use from AI workloads, these concerns are not accompanied by clear proposals for enforceable environmental limits or minimum performance standards. Moreover, there is little consideration of whether Europe's digital infrastructure needs can be met more efficiently rather than simply more expansively.

ECOS therefore calls for the Act to incorporate sufficiency principles, address life-cycle impacts, and prioritise infrastructure that is not only high-performing but also materially efficient, climate-aligned, and publicly accountable. A public consultation was launched to inform the impact assessment with a deadline of June 2025.

Voluntary initiatives

Code of Conduct for Data Centres

The European Code of Conduct (CoC) for Data Centres, launched by the Joint Research Centre in 2008, is a voluntary initiative aimed at improving energy efficiency across data centre infrastructure. It encourages operators to assess their IT and facility loads, develop an energy-saving action plan, and track performance indicators over time — primarily Power Usage Effectiveness (PUE). While the scheme has attracted participation from over 290 data centres and more than 120 organisations, its impact is inherently limited.

Participants are free to withdraw at any time, and there are no penalties for failing to meet efficiency commitments. The Code of Conduct relies on self-reporting and goodwill, rather than standardised third-party verification, and offers public recognition — such as plaques, logo use, and awards — as its main incentive. While some participants have demonstrated measurable gains (e.g., PUE below 1.80), there is no mechanism to ensure consistent application or comparable benchmarks across countries or providers.

Climate Neutral Data Centre Pact

The Climate Neutral Data Centre Pact (CNDCP), launched in 2021 by industry associations including CISPE, The Voice of Cloud Infrastructure Service Providers in Europe, and EUDCA, The European Data Centre Association, is a self-regulatory initiative through which major cloud and colocation providers commit to achieving climate neutrality by 2030. The pact outlines voluntary targets on energy efficiency, 100% renewable energy sourcing, water conservation, waste heat reuse, and circular economy measures. The European Commission has formally acknowledged the pact as the implementation and monitoring mechanism for data centres under the European Green Deal and the EU Digital Strategy — effectively outsourcing sectoral oversight to an industry-led coalition.

However, this pact remains outside the scope of EU legislation, with no binding enforcement, no legal accountability, and no independent regulatory authority overseeing compliance. Reporting is self-managed by participants, and external verification is limited to agreed-upon internal mechanisms. Key metrics, such as “renewable energy use,” may be fulfilled through market-based instruments like Guarantees of Origin (GOs), without requiring physical additionality or proximity. While the initiative signals a degree of industry alignment with EU climate goals, its governance structure — controlled by the very actors it aims to regulate — raises concerns about credibility, transparency, and the risk of greenwashing.

Recommendations

Current regulatory measures on data centres in the EU represent a preliminary yet modest step in addressing their environmental impacts. Initiatives such as the Energy Efficiency Directive (EED) focus primarily on monitoring energy and resource usage, falling short of imposing binding targets on energy and water consumption. While energy efficiency is an important component of sustainability, it must be accompanied by limits on resource use to mitigate the significant environmental footprint of data centres.

Stricter frameworks that go beyond monitoring to incentivise reductions in overall energy and water use are necessary. Without such measures, the goal of operating within planetary boundaries will remain out of reach, and the environmental impacts of data centres will continue to grow. This is particularly concerning when considering upcoming initiatives such as the Cloud and AI Development Act, which proposes a rapid expansion of digital infrastructure without corresponding binding environmental safeguards.

To ensure data centres remain within planetary boundaries, we recommend to:

Focus on more than just efficiency

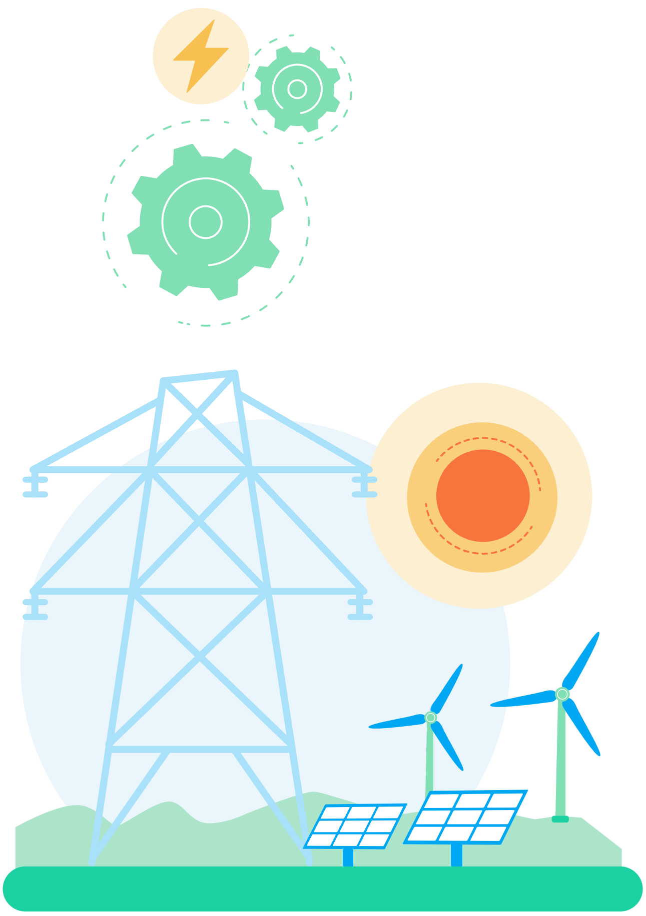
Improving the efficiency of data centres is a widely endorsed approach. Advances in energy-efficient hardware, optimised cooling systems, and better data management have reduced the per-unit energy consumption of digital services with relative success. However, overall energy demand continues to rise due to the exponential growth of data-driven applications.

This phenomenon reflects the "rebound effect," or Jevons paradox: as computing becomes more efficient and cheaper, its usage expands, ultimately negating environmental gains. For example, AI models are increasingly optimised to consume less energy per operation. Yet, this also enables more widespread use — e.g., for chatbots, personalised ads, or deep learning — which leads to higher aggregate consumption.

Efficiency therefore provides only a temporary fix, despite its undeniable merits. To address this, legally binding, **sector-specific targets should be introduced to reduce energy consumption and greenhouse gas emissions in data centres, aligned with EU climate goals**. These targets must also include measures to ensure the circularity of data centre infrastructure — from design to end-of-life. Public funding and incentives should be conditional on companies demonstrating meaningful progress toward environmental and social goals, rather than simply favouring technologies or sectors.

Furthermore, **data centres must be fully integrated into national energy strategies to avoid overburdening electricity grids and prevent fossil fuel lock-in**. Data centres must be required to source 100% renewable electricity, with strict additionality criteria to ensure that only newly added renewable capacity counts — an essential condition for the genuine phase-out of fossil fuels.

Finally, it is crucial that the data centre and AI boom is re-evaluated regarding its actual, non-artificially inflated demand by the market, and its real purpose and contribution to society.



Prioritise transparency

Transparency and robust data on environmental impacts (energy, water, emissions, materials) are essential for corporate accountability and informed decision-making.

ECOS strongly supports and advocates for the development of harmonised, science-based emissions calculation methodologies and life-cycle assessment frameworks for data centres and AI in standards, as these provide the technical foundation necessary to shape credible and coherent sustainability practices across the sector.

Without a common standardisation framework, environmental requirements risk being inconsistently applied, undermining both ambition and accountability. Standardisation offers a vital opportunity to define best practices, embed environmental integrity into digital infrastructure, and support broader climate and resource-efficiency objectives.

Environmental reporting remains fragmented and inconsistent, with tools varying widely in scope — from energy use during operation to embedded manufacturing emissions — making comparisons difficult and often misleading. By ensuring transparency, accuracy, and comparability in environmental data, sustainability claims made by data centre operators can be effectively put to the test and prevent greenwashing. Guarantees of origin (GOs), Renewable energy certificates (RECs) and carbon credits should not be an option.However, data alone cannot address environmental challenges. Even with improved methodologies, the structural drivers of overconsumption and unchecked infrastructure expansion must be tackled in parallel. Treating these issues as a data problem can lead to overlooking the root cause: the scale and speed of digital infrastructure growth. Delaying systemic action in pursuit of perfect data fosters inaction.

Operators already have enough information to take responsibility for their environmental footprint — what is missing are binding limits and meaningful structural changes.

Adopt sufficiency principles

Re-evaluating the push for more AI-driven data centre infrastructure entails re-examining how much of it we really need and whether we can make do without it — an approach that rests on principles of sufficiency.

Sufficiency aims at reducing absolute resource consumption by questioning the demand for goods and services.[44] It differs fundamentally from efficiency, which uses technological advancements to optimise resource use and lower input while still delivering the same output. Instead, sufficiency seeks to reduce unnecessary demand by encouraging changes in behaviour, policy, and design that align consumption with fundamental human and societal needs.

This concept is particularly relevant in the context of AI and data centres, where the relentless pursuit of increased computational power and storage capacity often results in unnecessary technological excess. By adopting a sufficiency-driven approach, policymakers and businesses can prioritise essential digital services, discourage wasteful consumption, and create frameworks that support responsible resource use. More sufficiency would not mean that efficiency should be abandoned, but that it should be complemented.

By adopting sufficiency principles, the digital sector can align with planetary boundaries, reducing both direct and indirect emissions while still delivering critical services that enhance human well-being.

A sufficiency approach could promote measures that limit technological excess by:

- **Prioritising essential digital services** — such as healthcare, education, science, and public welfare — while discouraging non-essential AI applications.
- **Optimising existing data centre infrastructure to respect planetary boundaries** by ensuring current facilities are used efficiently before permitting new ones, and requiring justification for expansion to avoid grid stress and unnecessary resource use.
- **Managing digital workloads and data storage more soberly**, for example through shared infrastructure and efforts to reduce superfluous or duplicated data.
- **Raising awareness among consumers and businesses** about the environmental impacts of digital consumption and promoting more conscious use.

Differentiate between essential and non-essential uses of AI and data centre services

A crucial step in implementing sufficiency within the digital sector is distinguishing between essential and non-essential uses of AI and data centre services. This differentiation would help to focus resources where they are most needed, preventing the indiscriminate expansion of digital technologies.

In response to the European Commission's consultations, the Shift Project, a French nonprofit think tank, proposed a framework to address sustainability concerns within the industry.[45] The proposal called for the creation of a service taxonomy to categorise IT services according to their societal value, distinguishing between essential public services, critical business operations, important industry services, and nice-to-have services. Their taxonomy aims to improve societal resilience during periods of high electricity demand by allowing some IT services to be safely reduced to save energy for more critical needs. By implementing such a taxonomy, data centre operators can identify and stop non-essential services, efficiently reducing energy consumption without shifting to more carbon-intensive power supplies. This approach would ensure that critical areas are prioritised during energy shortages.

Non-essential uses of AI and data centre services are those that primarily serve commercial interests or provide marginal improvements to convenience but come with disproportionate environmental costs.

For example:

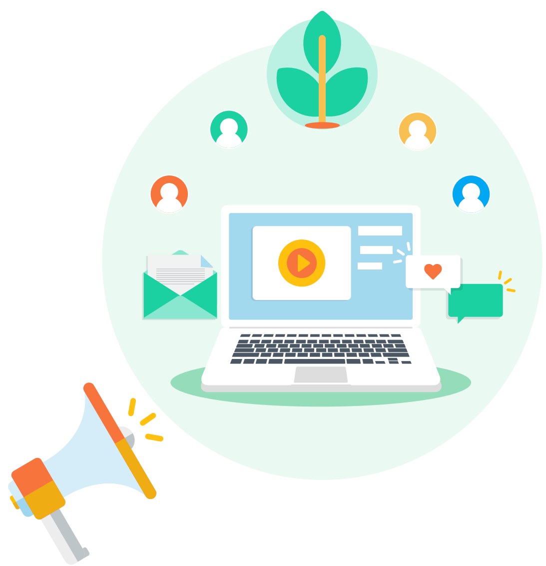
- **Personalised advertising:** AI systems that profile users and deliver targeted ads are energy-intensive and contribute significantly to data centre workloads, with limited societal value beyond advertising effectiveness.
- **Entertainment and social media algorithms:** Recommendation systems designed to maximise screen time continuously process and personalise content, driving unnecessary resource use with limited public benefit.
- **High-frequency trading:** AI-driven financial trading exploits millisecond advantages and requires energy-intensive colocation in financial data centres, benefiting a handful of firms while contributing little to wider economic stability.
- **Consumer-facing novelty applications:** Chatbots for trivial queries, virtual influencers, cosmetic filters, and other “AI for entertainment” services often rely on cloud-based processing, adding to data centre demand without delivering essential societal value.
- **AI for mass content creation:** Generative AI tools used to automatically produce large volumes of synthetic images, videos, or written content consume substantial computational power. While some creative applications could have educational value (e.g., in statistics or science), much of this output serves commercial marketing or promotional purposes, raising questions about whether the environmental costs are justified.

By identifying and reducing the emphasis on non-essential applications, businesses and policymakers can implement sufficiency-driven strategies that optimise digital infrastructure for meaningful outcomes.

Give society a choice

To ensure digitalisation aligns with societal interests, it is essential to enable informed participation and oversight. Consumer-facing campaigns should be developed to raise awareness of the environmental impacts associated with digital services and to promote more sustainable digital behaviours. Where appropriate, product labelling schemes for AI and digital services can support more conscious choices.

In parallel, local communities must be actively involved in decisions regarding new data centre infrastructure. Operators should engage stakeholders from the earliest planning stages, conduct comprehensive consultations and impact assessments across the supply chain, and prioritise inclusive processes. Particular attention should be given to the meaningful inclusion of marginalised and vulnerable groups to ensure that concerns related to environmental burdens, energy use, and social implications are adequately addressed.



Annex I
Types of data centres

Data centres are used to facilitate many different types of digital activities — there is no such thing as 'one size fits all'. The most prevalent categories of data centre are:

- **Enterprise (on-premises) data centres** are owned and operated by the same company that uses the infrastructure. All servers, storage, and networking gear reside on the company's own premises.

- **Colocation facilities** are third-party data centres where multiple organisations share the same facility. Instead of building their own, a company can co-locate its servers and equipment in a vendor's data centre space. In a traditional colocation model, the client owns the servers and storage hardware but rents physical space (racks or cages) along with power and network connectivity from the facility. The colocation provider handles the building infrastructure — such as power delivery, cooling, fire suppression, physical security, and internet connectivity — while the client manages and maintains their own equipment.

- **Hyperscale data centres** are massive facilities designed for scalability and high throughput, typically run by major cloud providers or tech giants. Companies like Amazon Web Services, Google, Microsoft Azure, and IBM operate hyperscale data centres that may span hundreds of thousands or even millions of square feet. These facilities contain tens of thousands of servers and leverage extensive automation. Hyperscale centres are built to serve millions of users and can handle a wide range of services — from cloud storage and computing to machine learning APIs — often for multiple customers at once.

  For example, a single hyperscale data centre might have at least 5,000 servers and enormous networks interconnecting them. They are engineered for efficiency and resilience, with redundant systems and sophisticated software to distribute workloads globally. Hyperscale data centres underpin public cloud infrastructure: when a company "moves to the cloud," its applications end up running in these huge, shared facilities. Notably, hyperscale data centres also deploy smaller regional installations to improve performance — known as **edge data centres**.

- **Edge data centres** are small, distributed facilities located closer to end-users or data sources. Unlike a centralised cloud data centre that might be hundreds of miles away, an edge data centre could be in the same city or even on the premises where data is generated. The goal is to reduce latency (delay) by processing data locally rather than sending it across long distances. Edge data centres often support real-time applications like streaming, online gaming, industrial Internet of Things (IoT), or autonomous systems. For instance, telecom providers might place edge servers at base stations or central offices to quickly handle mobile AI services or content delivery for local users. These sites are typically much smaller than corporate or hyperscale facilities, but they still include the core components (servers, storage, networking, cooling) in a compact form.

Notes and references

1 https://www.iea.org/reports/global-energy-review-2025

2 In February 2025, the European Commission announced at the AI Summit in Paris that it would mobilise EUR 200 billion from public and private funds for AI infrastructure, of which EUR 20 billion will be allocated to AI gigafactories. Earlier, the Commission had announced an initiative for seven AI factories in seven different EU Member States. European governments have also announced investments or plans to support the development of AI infrastructure. For example, France introduced a EUR 109 billion investment plan including identifying 35 sites around France for new data centres and a 1GW data centre plan with the United Arab Emirates (UAE).

3 The European Court of Auditors' recent criticism of the EU Chips Act offers a cautionary parallel for AI and data centre policy. The auditors concluded that the EU's microchip strategy is "deeply disconnected from reality," citing fragmented funding, unrealistic targets, and lack of coordination. These same structural flaws are evident in the EU's approach to AI infrastructure: political momentum is prioritising scale and speed without a solid foundation in environmental governance, system readiness, or critical reflection on societal value.

4 https://ecostandard.org/news_events/climate-neutral-clouds-the-journey-towards-sustainable-and-energy-efficient-data-centres-new-briefing/

5 https://www.osc.edu/press/ohio_supercomputer_center_launches_gpu_cluster_ascend_for_high_performance_computing_work

6 https://www.tomshardware.com/news/meta-supercomputer-16000-a100-gpus

7 https://www.iea.org/reports/global-energy-review-2025

8 https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

9 https://www.iea.org/commentaries/what-the-data-centre-and-ai-boom-could-mean-for-the-energy-sector

10 https://publications.jrc.ec.europa.eu/repository/bitstream/JRC135926/JRC135926_01.pdf

11 https://www.reuters.com/technology/europes-data-centre-power-demand-expected-triple-by-2030-mckinsey-report-says-2024-10-23

12 https://cc-techgroup.com/data-center-energy-consumption/

13 Furthermore, Bitcoin mining has a significant water footprint, with studies showing that between January 2020 and December 2021, its water usage was equivalent to 660,000 Olympic-sized swimming pools. Source: https://www.polytechnique-insights.com/en/columns/energy/bitcoin-electricity-consumption-comparable-to-that-of-poland/

14 https://carboncredits.com/u-s-data-centers-power-demand-surges-to-46000-mw-whats-driving-the-growth/

15 https://greengaming.lbl.gov/cloud-gaming

16 https://www.theguardian.com/technology/article/2024/jul/02/google-ai-emissions

17 https://trellis.net/article/microsoft-pivots-climate-investments/

18 https://trellis.net/article/the-meta-dilemma-invest-billions-in-ai-but-find-ways-to-cut-emissions-too/

19 https://www.geekwire.com/2024/amazons-carbon-footprint-shrinks-3-last-year-but-ai-driven-climate-challenges-loom/

20 https://www.wsj.com/business/energy-oil/oil-majors-flirt-with-electricity-0d1df707

21 Ibid.

22 https://www.energymonitor.ai/sectors/power/small-modular-reactors-smrs-what-is-taking-so-long/

23 https://www.datacenterdynamics.com/en/news/exxonmobil-plots-natural-gas-power-plant-to-exclusively-power-data-centers

24 https://www.reuters.com/business/energy/chevron-working-supply-power-data-centers-executive-says-2024-12-11

25 https://www.belganewsagency.eu/flemish-power-grid-not-strong-enough-to-fully-connect-all-businesses

26 https://www.theguardian.com/technology/2024/sep/15/data-center-gas-emissions-tech

27 https://www.bloomberg.com/news/articles/2024-08-21/ai-tech-giants-hide-dirty-energy-with-outdated-carbon-accounting-rules

28 https://ecostandard.org/publications/climate-neutrality-report/

29 https://www.ft.com/content/3ac7e92f-bea5-4460-8791-23a2697d8b80

30 https://www.theguardian.com/environment/2023/sep/19/do-carbon-credit-reduce-emissions-greenhouse-gases

31 https://www.staxengineering.com/stax-hub/the-environmental-impact-of-data-centers/

32 https://www.savills.com/research_articles/255800/345047-0

33 https://www.ft.com/content/1d468bd2-6712-4cdd-ac71-21e0ace2d048

34 https://www.thetimes.com/uk/technology-uk/article/thirsty-chatgpt-uses-four-times-more-water-than-previously-thought-bc0pqswdr

35 https://www.itu.int/en/ITU-D/Environment/Documents/Toolbox/GEM_2020_def.pdf

36 https://www.reuters.com/markets/commodities/ai-could-add-1-million-tons-copper-demand-by-2030-says-trafigura-2024-04-08/

37 https://www.theguardian.com/environment/2024/oct/16/global-water-crisis-food-production-at-risk

38 https://www.theguardian.com/environment/2024/sep/15/data-centers-energy-consumption

39 https://www.theguardian.com/global-development/2024/sep/25/mexico-datacentre-amazon-google-queretaro-water-electricity

40 UNEP. (2024). AI Environmental Impact Issues Note. United Nations Environment Programme: https://wedocs.unep.org/handle/20.500.11822/46288;jsessionid=06C4DCB996D913FA1A487CDB4B9DDA0A

41 Ibid.

42 https://apnews.com/article/lithium-mining-energy-transition-child-labor-nigeria-africa-22155590dddf7ecc0b9fd55b221c6d9f

43 https://data.consilium.europa.eu/doc/document/ST-6775-2020-INIT/en/pdf

44 https://ecostandard.org/wp-content/uploads/2024/04/ECOS-2024-CMT-044-Sufficiency_manifesto.pdf

45 https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13818-Data-centres-in-Europe-reporting-scheme/F3451089_en

Environmental Coalition on Standards

c/o WeWork
Rue du Commerce 31
1000 Brussels, Belgium
+32 2 899 76 80
ecostandard.org

Follow us

@ecos.ngo

ECOS-NGO