THE OXFORD INSTITUTE FOR ENERGY STUDIES

May 2025: ISSUE 145

forum

A QUARTERLY JOURNAL FOR DEBATING ENERGY ISSUES AND POLICIES

ARTIFICIAL INTELLIGENCE AND ITS IMPLICATIONS FOR ELECTRICITY SYSTEMS

Third, the geopolitical race to expand AI investment risks delaying the transition to clean power, further entrenching dependence on fossil fuels. It could also conflict with other policies, notably affordable electricity prices. Ironically, given the interest in environmental sustainability, the economic risks are substantial with renewable energies, which are capital-intensive. Because these assets are intermittent, they will require other capital-intensive resources to back them up, including storage and networks. If these assets are more than needed and require government subsidies, the costs will be borne by taxpayers or electricity consumers. To mitigate these risks, electricity investments should respond to competitive market signals rather than geopolitical competition, central planning, and subsidies.

Finally, given the economic, environmental, and societal risks of global expansion of AI data centres, there is a need for the sorts of actions identified in the AI Action Summit statement. Concretely, we need to promote multistakeholder dialogues on AI and the environment, to study closely the energy impact of AI, and to reach global agreements on the use of sustainable energy for AI data centres.

AI'S INDIRECT IMPACTS ON CLIMATE OUTWEIGH CONCERNS OVER ITS DIRECT ENERGY FOOTPRINT

Charlie Wilson, Yee Van Fan, and Felippa Amanta

Weekly headlines on AI data centres herald soaring energy footprints, eye-catching contracts for low-carbon power, alarm over electricity network congestion, and backsliding of tech companies' net-zero commitments. Amid this emphasis on the direct impact of AI on energy and greenhouse gas (GHG) emissions, the implications of how and for what AI is used are less discussed—at least in relation to energy and climate.

A simple taxonomy of AI's impacts on energy distinguishes direct, indirect, and systemic impacts.[1] Direct impacts are from the energy consumed by AI infrastructure like data centres. Indirect impacts are from the energy consumed or saved by the AI applications that this infrastructure makes possible. As AI is a general-purpose technology, these indirect impacts occur throughout almost all forms of economic and social activity. Systemic impacts are harder to isolate and quantify but include the implications for energy demand of the structural changes wrought by AI on economic systems (e.g. from industrial to service economies) and on social systems (e.g. from physical to virtual modes of interaction).

This article sets current debates around AI's direct energy impact in the context of evidence on its indirect impacts. It also briefly discusses the challenges and opportunities for AI governance to mitigate environmental risks. It argues that the direct energy impact of AI is problematic locally rather than globally, and that the indirect energy impact of AI is larger, more uncertain, more diffuse, and harder to regulate—and so of greater concern.

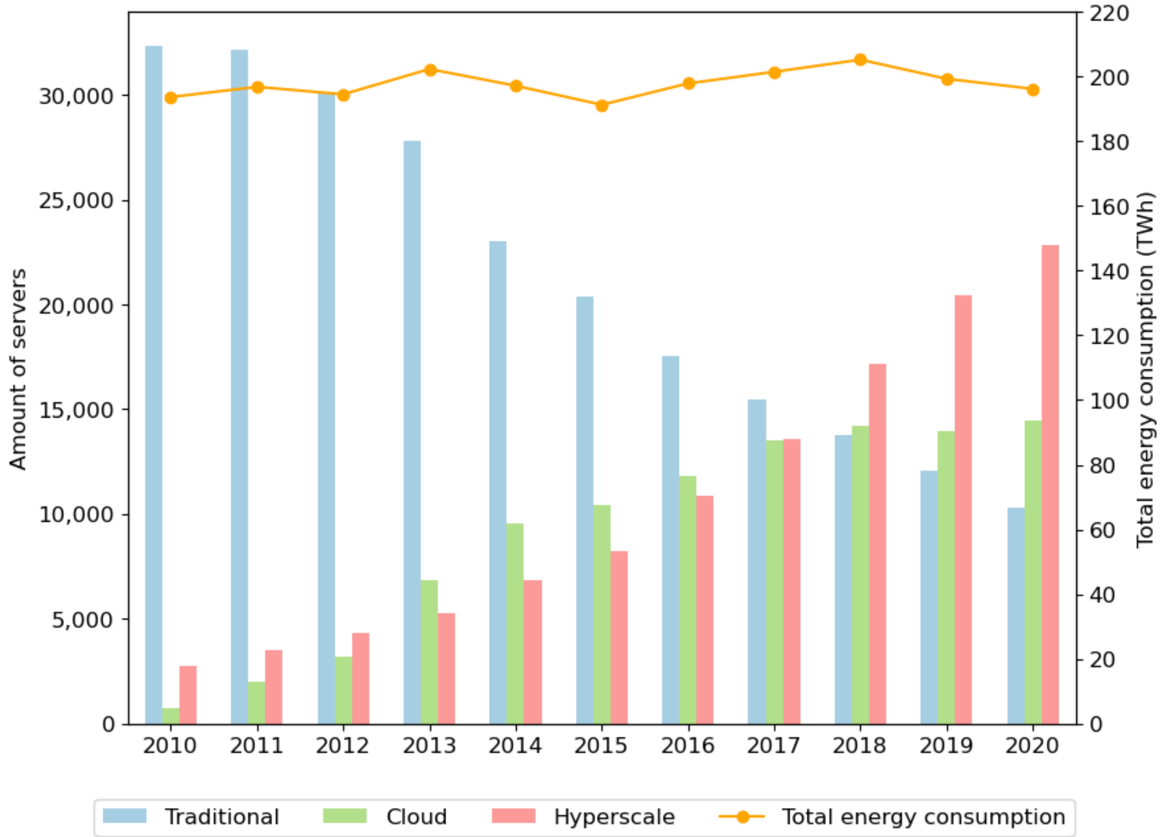
Direct energy impacts: current issues and near-term projections

Globally, information and communication technology (ICT) infrastructure accounts for around 2–4 per cent of total electricity consumption and a similar share of GHG emissions.[2] It varies by country: around 5 per cent in the US, 4 per cent in the EU, and 3 per cent in China.[3] These totals break down very roughly one-third each between data centres, networks, and end-use devices, but the data centre share is increasing.

The decade to 2020 saw exponential increases in the demand for computation driven by pervasive uptake of cloud computing. Despite this, data centre electricity consumption remained broadly flat (Figure 1).[4] This was achieved through a combination of rapid technological improvements in the energy efficiency of data centre servers and infrastructure (e.g. cooling systems), and operational economies of scale as data centres jumped in size from conventional to cloud to hyper scales.

[1] Kaack, L. H., P. L. Donti, E. Strubell, G. Kamiya, F. Creutzig, and D. Rolnick (2022), 'Aligning artificial intelligence with climate change mitigation', *Nature Climate Change*, https://doi.org/10.1038/s41558-022-01377-7.
[2] Freitag, C., M. Berners-Lee, K. Widdicks, B. Knowles, G. S. Blair, and A. Friday (2022), 'The real climate and transformative impact of ICT: a critique of estimates, trends and regulations', *Patterns* 2(9), 100340, https://doi.org/10.1016/j.patter.2021.100340.
[3] International Energy Agency (2024), *Electricity 2024: Analysis and Forecast to 2026*, Paris: IEA, https://www.iea.org/reports/electricity-2024
[4] Masanet, E., A. Shehabi, N. Lei, S. Smith, and J. Koomey (2020), 'Recalibrating global data center energy-use estimates', *Science* 367(6481), 984, https://doi.org/10.1126/science.aba3758.

Figure 1: Historical energy consumption of data centres globally



Source: Masanet, E., A. Shehabi, N. Lei, S. Smith, and J. Koomey (2020), 'Recalibrating global data center energy-use estimates', *Science* 367(6481): 984, https://doi.org/10.1126/science.aba3758

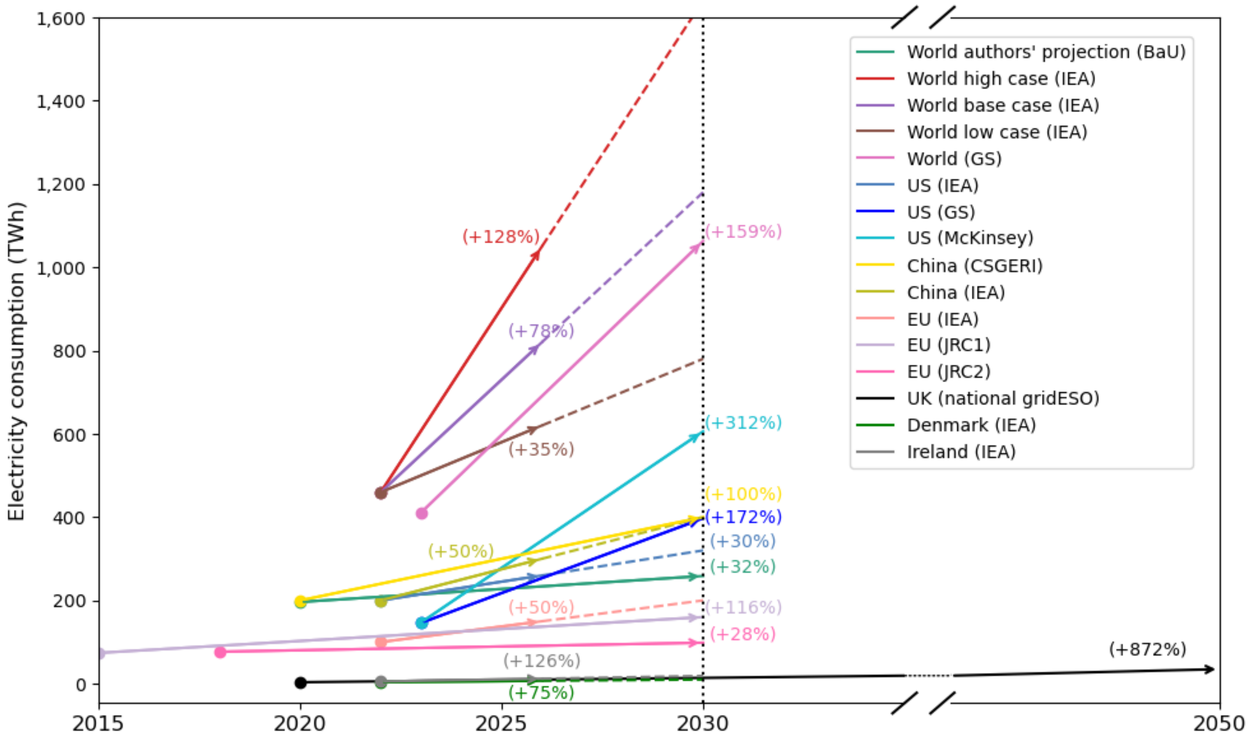
This market-led innovation response to increasing demand for computation associated historically with cloud computing is instructive for the current situation with energy-hungry training and inference of generative AI (genAI) models.

The increasing scale and widespread use of genAI models since the launch of ChatGPT's chatbot in November 2022 have driven up computational demands on data centres. Some near-term projections estimate three-, six-, and even eight-fold increases in electricity consumption as a result over the next three to five years (Figure 2). Older projections made before the genAI boom show much slower increases.

For comparison purposes, we fitted a statistical model to data centre electricity consumption over the period 2010–2021 (Figure 1). Four variables explained observed variation over time and world regions: GDP, population, trade, and climate (cooling degree days). We then combined our model with future projections of these explanatory variables[5] to estimate future data centre electricity demand ('authors' projection' in Figure 2). It is more conservative than others as it is calibrated to the historical period of data centre efficiency improvements and hyperscaling.

[5] SSP Database (Shared Socioeconomic Pathways) – Version 2, December 2018, Laxenburg, Austria: IIASA, https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about

Figure 2: Projected increase in data centre electricity consumption in different regions



Source: Various

Notes: Circle markers show level in year projection was made; arrowheads show level in year projection ends; dotted lines indicate linear extrapolation to 2030. BaU = Business as Usual. IEA = International Energy Agency[6]; GS = Goldman Sachs[7]; McKinsey = McKinsey & Company[8], CSGERI = China's State Grid Energy Research Institute[9], JRC = Joint Research Centre[10]; ESO = Energy System Operator[11]

Whether a similar innovation response could dampen skyrocketing growth in data centre energy needs depends on several factors including demand for computation from genAI, innovation to develop more efficient AI models, and efficiency gains in data centre hardware, software, and infrastructure.

First, will compelling AI use cases become widely adopted? GenAI holds significant promise—an early McKinsey estimate put this at $2.6–4.4 trillion globally.[12] AI chatbots have become widely used both by service providers and consumers,[13] driving up data centre activity. However, early-stage markets for new technologies are often associated with hype. An eventual 'shakeout' of AI use cases will leave only those with compelling benefits that justify the computational and energy costs.

Second, will AI models become more energy efficient? Energy accounts for about 20 per cent of the cost base for data centre business models; ICT hardware is the dominant cost driver.[14] Since the initial 'bigger is better' phase of model development,

[6] International Energy Agency (2024), 'Electricity 2024: Analysis and Forecast to 2026', Paris: IEA, https://www.iea.org/reports/electricity-2024

[7] Goldman Sachs (2024), 'AI is poised to drive 160% increase in data center power demand', https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand

[8] McKinsey (2024), 'Data centers and AI: How the energy sector can meet power demand', https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power

[9] China Daily (2021), 'Green data centers in focus', https://english.www.gov.cn/statecouncil/ministries/202112/09/content_WS61b13edac6d09c94e48a1f81.html

[10] Kamiya, G. and Bertoldi, P (2024), 'Energy consumption in data centres and broadband communication networks in the EU', EC Joint Research Centre (JRC), https://op.europa.eu/en/publication-detail/-/publication/f6822d03-cedb-11ee-b9d9-01aa75ed71a1/language-en#

[11] nationalgridESO (2022), 'Data Centres What Are Data Centres and How Will They Influence the Future Energy System? https://www.neso.energy/document/246446/download.

[12] McKinsey & Company (2024), 'How data centres and the energy sector can sate AI's hunger for power', San Francisco: McKinsey, https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power

[13] Purdy, M. (2024),'What Is Agentic AI, and How Will It Change Work?', Harvard Business Review, https://hbr.org/2024/12/what-is-agentic-ai-and-how-will-it-change-work.

ChatGPT Statistics (March 2025), 'Number of Users & Queries', https://www.demandsage.com/chatgpt-statistics/

[14] McKinsey & Company (2024), 'How data centres and the energy sector can sate AI's hunger for power', San Francisco: McKinsey, https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power

genAI developers have already been moving towards leaner code and smaller, task-specific models.[15] The recent launch of DeepSeek's R1, which outperforms larger genAI models yet requires far fewer high-performance GPU chips to run, further evidenced the potential for both computational and energy efficiency improvements.[16]

Third, will data centres become more energy efficient? Koomey's Law is the energy-efficiency analogue of the better known Moore's Law, which charts exponential improvements in computational power observed in semiconductors.[17] Koomey's Law describes a similar trajectory for the energy required per unit of computation. It helps explain the flat energy consumption trends in data centres historically (Figure 1). However, it was observed originally from the 1950s up to the early 2010s. Whether it can continue through new AI chip architectures and less software bloat is an area of debate and innovation, but signs are positive.[18] Google and Nvidia have reported 80-fold and 25-fold improvements in the energy performance of their new AI chips, respectively. The design and scale of data centres also continues to improve, with best-in-class examples making use of 'free' cooling from natural heat sinks that reduce energy consumption for non-ICT functions to a minimum.[19]

A fourth question for GHG emissions is: will sufficient and additional low-carbon power be available to meet data centre needs? Tech companies have strong GHG emission-reduction commitments and deep pockets and are relatively price insensitive when it comes to electricity. These three criteria could spur a nuclear renaissance. Google and Oracle have reportedly ordered a series of small modular nuclear reactors (SMRs) to power their data centres.[20] The UK government has also just announced a commitment to both hyperscale data centres and SMRs.[21] However, there is a mismatch in time frames between the time needed for data centre permitting and construction, and the much longer time needed to permit and build new nuclear capacity including SMRs which remain unproven.

In contrast, large-scale wind or solar plants can be built in similar time frames as data centres, absent supply chain constraints. Data centres are already major buyers of renewable generation. Whether this drives additional investment depends on other constraints facing renewables, particularly grid connections, and the need for back-up, storage, or flexible demand to balance intermittent output. These constraints vary by location and are already starting to exert more influence on where data centres are built.

Taking all these factors into account, it is highly likely that global data centre energy use will rise, due to genAI's demand for computation, but not by as much as some of the more exuberant projections suggest, due to the efficiency response across all levels from model development to infrastructure.

The GHG emission consequences of this are important but manageable—electricity is rapidly decarbonizing.

However, global averages mask important localized impacts. Problems with grid capacity and congestion have led to de facto moratoria on new data centres in the Netherlands, Singapore, and elsewhere.[22] Localized impacts will depend on whether location decisions are driven by computational or energy logics.

Locating data centres near cities or industrial areas reduces latency in data transmission and reduces the need for new ICT network infrastructure.[23] Locating data centres in areas with sufficient electricity network capacity reduces permitting and connection delays. Locating them in areas with readily available heat sinks and low-carbon electricity reduces cooling energy needs and GHG emissions, respectively.[24]

---

[15] The Economist (2025), 'OpenAI's latest model will change the economics of software', https://www.economist.com/business/2025/01/20/openais-latest-model-will-change-the-economics-of-software.
[16] NBC News (2025), 'Why DeepSeek is different, in three charts', https://www.nbcnews.com/data-graphics/deepseek-ai-comparison-openai-chatgpt-google-gemini-meta-llama-rcna189568.
[17] Koomey, J., S. Berard, M. Sanchez, and H. Wong (2011), 'Implications of historical trends in the electrical efficiency of computing', *IEEE Annals of the History of Computing* 33(3), 46–54, https://doi.org/10.1109/MAHC.2010.28.
[18]IEA (2025), 'Efficiency improvement of AI related computer chips, 2008-2023', https://www.iea.org/data-and-statistics/charts/efficiency-improvement-of-ai-related-computer-chips-2008-2023
Prieto, A., B. Prieto, J. J. Escobar, and T. Lampert (2024), 'Evolution of computing energy efficiency: Koomey's law revisited', *Cluster Computing* 28(1), 42, https://doi.org/10.1007/s10586-024-04767-y.
[19] Climate Neutral Data Centre Pact (2025), https://www.climateneutraldatacentre.net.
[20] Financial Times (2025), 'Google orders small modular nuclear reactors for its data centres', https://www.ft.com/content/29eaf03f-4970-40da-ae7c-c8b3283069da.
[21] World Nuclear News (2025), 'UK government considering role for SMRs in AI expansion', https://www.world-nuclear-news.org/articles/uk-government-considering-role-for-smrs-in-ai-expansion.
[22] International Energy Agency (2024), *Electricity 2024: Analysis and Forecast to 2026*, Paris: IEA.
[23] Oxford Institute for Energy Studies (2024), *Global Electricity Demand: What's Driving Growth and Why It Matters?* Oxford: OIES.
[24] McKinsey & Company (2024), 'How data centres and the energy sector can sate AI's hunger for power', San Francisco: McKinsey, https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power.

Indirect energy impacts: AI applications across sectors

The magnitude of AI's indirect impacts is considerably larger than its direct energy footprint.[25] Indirect impacts are the net outcome of energy-saving efficiency, substitution, and optimization effects from AI applications offset by energy-increasing rebound, induced demand, and intensification effects.

In other words, AI applications are a double-edged sword for energy. By reducing the time, cost, effort, or friction of a wide range of activities, AI drives up or induces demand for those activities (in addition to creating new classes of energy-intensive activity like cryptocurrency or agentic AI).

On which side the double-edged sword falls will vary for each AI application and its deployment context.

Some AI applications are unambiguously beneficial for energy and GHG emissions. Balancing supply and demand in real time on electricity networks with intermittent renewable generation and distributed storage and flexibility assets is increasingly complex, fast, and data-intensive: hallmarks of an AI use case. There are numerous examples of AI models being used to improve power system optimization, scheduling, and dispatch in support of the low-carbon transition.[26]

However, some AI applications are unambiguously detrimental for GHG emissions. For example, AI models are widely used in the oil and gas industry to improve efficiency, increase yield, and lower the cost of fossil fuel extraction.[27]

Most AI applications sit on neither of these unambiguous poles. Whether their use reduces energy or GHG emissions depends on rebound effects and whether they can be managed.

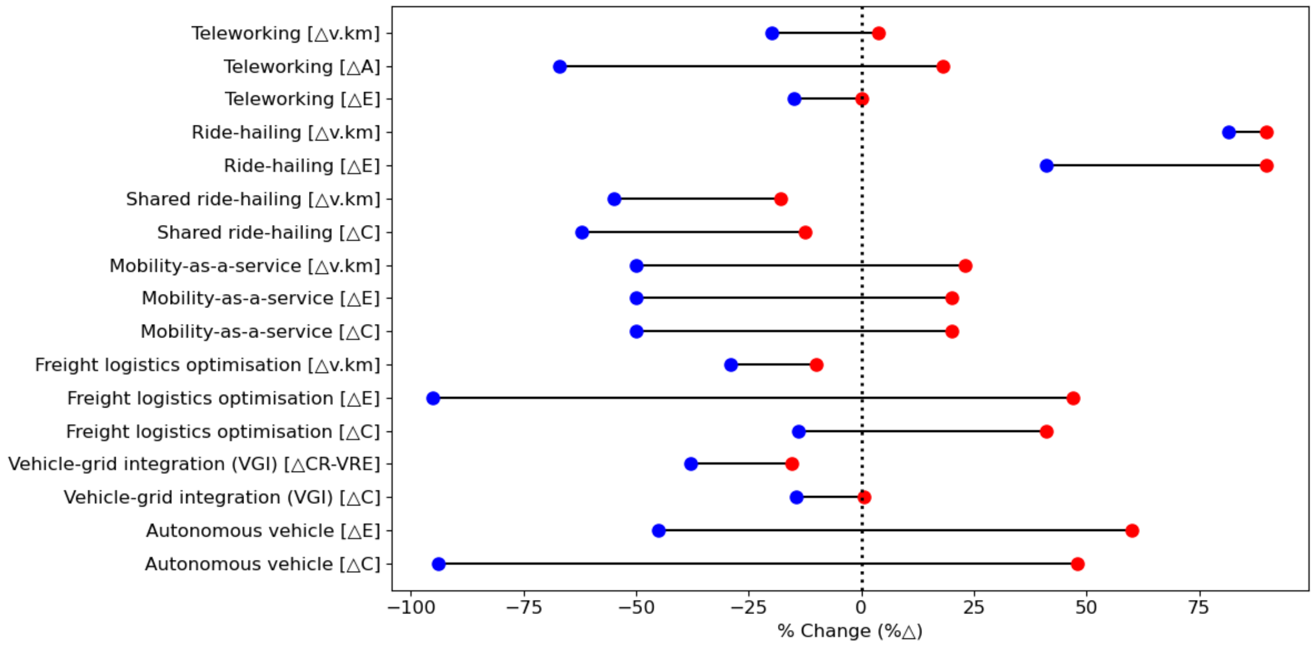
For example, digital twins and smart control systems help optimize the energy performance of industrial processes, urban traffic flows, and buildings. But if the resulting efficiency or productivity improvements lead to increased demand for industrial output, travel, or heating, then the net effect on energy becomes ambiguous—determined on a case-by-case basis.

The same applies to AI and now genAI shaping consumer preferences on retail platforms and through influencers.[28] Clothing and fashion exemplify both sides of the double-edged sword: reuse platforms like Vinted and fast fashion retailers like Shein have both seen phenomenal recent market growth.

We synthesized evidence on the indirect impacts of numerous digital (including AI) applications in the buildings, transport, and industrial sectors. Figure 3 gives an example for transport-related applications. Most applications have indirect impacts that can both reduce and increase energy use. Autonomous vehicles are an extreme example with the potential to halve or to double the energy consumed by urban vehicle fleets depending on the extent to which the savings from optimized vehicle routing, driving, and occupancy are offset by increases in vehicle trips and distances.[29]

[25] Scientific American (2023), 'AI's Climate Impact Goes beyond Its Emissions', https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions/.
[26] Kaack, L. H., P. L. Donti, E. Strubell, G. Kamiya, F. Creutzig, and D. Rolnick (2022), 'Aligning artificial intelligence with climate change mitigation', *Nature Climate Change*, https://doi.org/10.1038/s41558-022-01377-7.
[27] Tariq, Z., M. S. Aljawad, A. Hasan, M. Murtaza, E. Mohammed, A. El-Husseiny, S. A. Alarifi, M. Mahmoud, and A. Abdulraheem (2021), 'A systematic review of data science and machine learning applications to the oil and gas industry', *Journal of Petroleum Exploration and Production Technology* 11(12), 4339–4374, https://doi.org/10.1007/s13202-021-01302-2.
[28] Forbes (2023), 'This AI-Generated Influencer Can Pull In Almost $11,000 A Month', https://www.forbes.com/sites/lesliekatz/2023/11/24/this-ai-generated-influencer-can-pull-in-10000-euros-a-month/.
[29] Wadud, Z., D. MacKenzie, and P. Leiby (2016), 'Help or hindrance? the travel, energy and carbon impacts of highly automated vehicles', *Transportation Research Part A: Policy and Practice* 86, 1–18, https://doi.org/10.1016/j.tra.2015.12.001.

Figure 3: Empirical evidence for transport-related digital applications' impact on activity, energy, and GHGs



Notes: ΔA or Δv.km = change in activity or vehicle kilometres travelled; ΔE = change in energy use; ΔC = change in carbon emissions; ΔCR-VRE = change in curtailment rate of variable renewable energy.
Source: Wilson, C., et al. (2024), *Evidence Synthesis of Indirect Impacts of Digitalisation on Energy and Emissions*, 2024 10th International Conference on ICT for Sustainability, https://doi.org/10.1109/ICT4S64576.2024.00021

In sum, the indirect impacts of AI on energy demand are potentially large, diffuse, and generally both positive and negative. This poses challenges for mitigation.

Policy and governance considerations: challenges and potential regulatory responses

The direct impacts of AI on energy and GHG emissions are detrimental and largely confined to the electricity sector (except for the material, water, and land-use footprints of ICT infrastructure). The indirect impacts of AI applications can be detrimental or beneficial and are spread throughout the economy. Policy responses to direct and indirect impacts are therefore very different.

Direct impacts can be addressed through measures such as voluntary or mandatory energy performance standards for data centres,[30] scope 2 (electricity-related) emissions reporting requirements for data centre operators and tech companies, public sector procurement policies, land use planning and permitting, and electricity network regulations governing connections for large new loads. Examples of all these policies have been implemented in the US, China, the EU, Singapore, and other data centre locations and will help limit increases in energy demand.[31]

Indirect impacts are currently left to energy or climate policies in the many different AI application domains. For example, urban planning, traffic regulation, and safety rules govern the deployment of autonomous vehicles. This in turn will shape whether autonomous vehicles improve the efficiency of intra-urban travel flows or induce new forms of travel demand. As Figure 3 shows, these outcomes have markedly different implications for energy demand.

Managing the risk that AI applications may drive up energy demand as a result of how they change behaviours, processes, or economic activities is not part of current thinking on AI and energy governance. AI risk taxonomies[32] and regulatory frameworks like the EU's AI Act do recognize environmental risks, but for GHG emissions these are limited to direct impacts only.

Should the indirect impacts of AI applications on energy and GHG emissions be considered a systemic risk to be mitigated by AI governance? If so, what might such an approach look like?

[30] Climate Neutral Data Centre Pact (2025), https://www.climateneutraldatacentre.net.
[31] Brocklehurst, F. (2024), 'Policy development on energy efficiency of data centres', IEA Energy Efficient End-Use Equipment Technology Collaboration Programme, https://www.iea-4e.org/wp-content/uploads/2024/02/Policy-development-on-energy-efficiency-of-data-centres-draft-final-report-v1.05.pdf.
[32] Slattery, Peter, Alexander K. Saeri, Emily A. C. Grundy, Jess Graham, Michael Noetel, Risto Uuk, James Dao, et al. 2024. 'The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence'. doi:10.48550/arXiv.2408.12622.

The UK's Royal Society and the open-source AI company Hugging Face have independently mooted the idea of a 'proportionality' test for new AI applications,[33] under which those providing weak societal value but requiring large amounts of energy should be subject to regulatory guidance or bans.

UNESCO's Recommendations on the Ethics of AI include the preference for data- and energy-efficient methods in order to mitigate both direct and indirect environmental impacts: 'where there are disproportionate negative impacts on the environment, AI should not be used.'[34]

There are few signs yet of this happening, nor regulatory appetite to make it happen. However, a necessary first step would be towards a more systematic tracking and reporting framework for the indirect impacts of AI on energy. The International Telecommunications Union has recently published emissions accounting standards that support moves in this direction.[35] More systematic reporting in turn enables investors to orient their portfolios towards net GHG-reducing AI applications. It also enables policymakers to differentially incentivize AI applications in domains where their net (direct and indirect) GHG impact is aligned with decarbonization goals.

Conclusion—summary of key arguments and future outlook

AI and genAI models are means not ends. Their application amplifies and accelerates trends under prevailing business incentives, market logics, and governance conditions both towards GHG emission reductions and in the contrary direction towards energy profligacy and GHG emissions growth.

Currently, the narrower issue of AI's direct impact on electricity demand and networks has occupied analysts and regulators' attention. Despite steep projections for near-term global growth, challenges and regulatory responses will be localized.

The broader issue of AI's indirect impact on energy and GHG emissions is much more challenging, but will only increase in importance as AI applications continue to exert a transformative effect on diffuse economic and social activities.

SHOULD AI CHANGE THE WAY WE THINK ABOUT THE ENERGY SYSTEM?

Sam Young

The energy system is just that—a system of interconnected parts. When looking at a system, most people find it easier to think about the component parts than the connections between them. Yet the connections often drive the behaviour of the system. A rainforest is not just trees and animals—it is also the food chains, water cycles, and symbiotic relationships that keep it going. The same is true of the energy system.

AI will increasingly connect parts of the energy system that have not been strongly connected before. This will change the way the system behaves—for the better in some cases, and for worse in others. To account for this, we need to update how we think about the energy system. This article outlines three ways AI will influence energy system complexity and proposes changes to the sector's mental models that may be required as a result.

AI means small things matter

The first connection that AI enables is between numerous small assets across the system. As the energy system becomes more decentralized, moving from large power stations to distributed generation and flexible demand, it is faced with the need to coordinate the control of millions of small assets. AI enables the aggregation of these small assets—e.g. electric vehicles (EVs) and heat pumps—so they behave in a coordinated manner. This allows operators to treat them conceptually as a single large asset and continue dispatching as they always have.

[33] The Royal Society (2020), Digital Technology and the Planet: Harnessing Computing to Achieve Net Zero, London: The Royal Society, https://royalsociety.org/news-resources/projects/digital-technology-and-the-planet/
Luccioni, Alexandra Sasha, Emma Strubell, and Kate Crawford. 2025. 'From Efficiency Gains to Rebound Effects: The Problem of Jevons' Paradox in AI's Polarized Environmental Debate'. doi:10.48550/arXiv.2501.16548.
[34] UNESCO (2025), 'Recommendation on the Ethics of Artificial Intelligence', Document SHS/BIO/REC-AIETHICS/2021, https://unesdoc.unesco.org/ark:/48223/pf0000380455.
[35] International Telecommunication Union (2022), 'Enabling the Net Zero transition: assessing how the use of information and communication technology solutions impact greenhouse gas emissions of other sectors, recommendation ITU-T L.1480 (12/2022), Geneva: ITU, https://www.itu.int/ITU-T/recommendations/rec.aspx?id=15030&lang=en.