The impacts of AI on energy and emissions.

Charlie Wilson, Yee Van Fan, and Felippa Amanta

13 Feb 2025, iDODDLE.org

iDODDLE

The Impacts of Digitalised Daily Life on Climate Change

Introduction.

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 from indirect and systemic.[1] Direct impacts are the energy consumed by AI infrastructure like data centres but also networks and end-use devices. Indirect impacts are the energy consumed or saved by the AI applications that this infrastructure make 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).

In this article, we set current debates around AI's direct energy impact in the context of evidence on its indirect impacts. We also briefly discuss the challenges and opportunities for AI governance to mitigate environmental risks.

---

[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. doi.org/10.1038/s41558-022-01377-7

Our main arguments are twofold. First, the direct energy impact of AI is problematic locally rather than globally. Second, 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 & near-term projections

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

The decade to 2020 saw exponential increases in the demand for computation. This was driven particularly by pervasive uptake of cloud computing.[ii] Despite the dramatic increases in data centre activity that resulted, electricity consumption remained broadly flat [Fig 1].[4] This was achieved through a combination of rapid technological improvements in the energy efficiency of data centre infrastructure, both ICT-related and balance of plant (e.g., cooling systems), and operational economies of scale as data centres jumped in size from conventional to cloud to hyper scales. Typical power utilisation effectiveness (PUE) which measures the ratio of total energy use to ICT energy use in data centres fell globally from well over 2 to 1.6 - 1.7 with improvements seen particularly in the period 2007-2013.[5]

---

[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] IEA (2024). Electricity 2024: Analysis and forecast to 2026. Paris, France, International Energy Agency.

[4] Masanet, E., A. Shehabi, N. Lei, S. Smith and J. Koomey (2020). "Recalibrating global data center energy-use estimates." Science 367(6481): 984. 10.1126/science.aba3758

[5] Lawrence A., 2019. Is PUE actually going up, Uptime Institute, https://journal.uptimeinstitute.com/is-pue-actually-going-up/

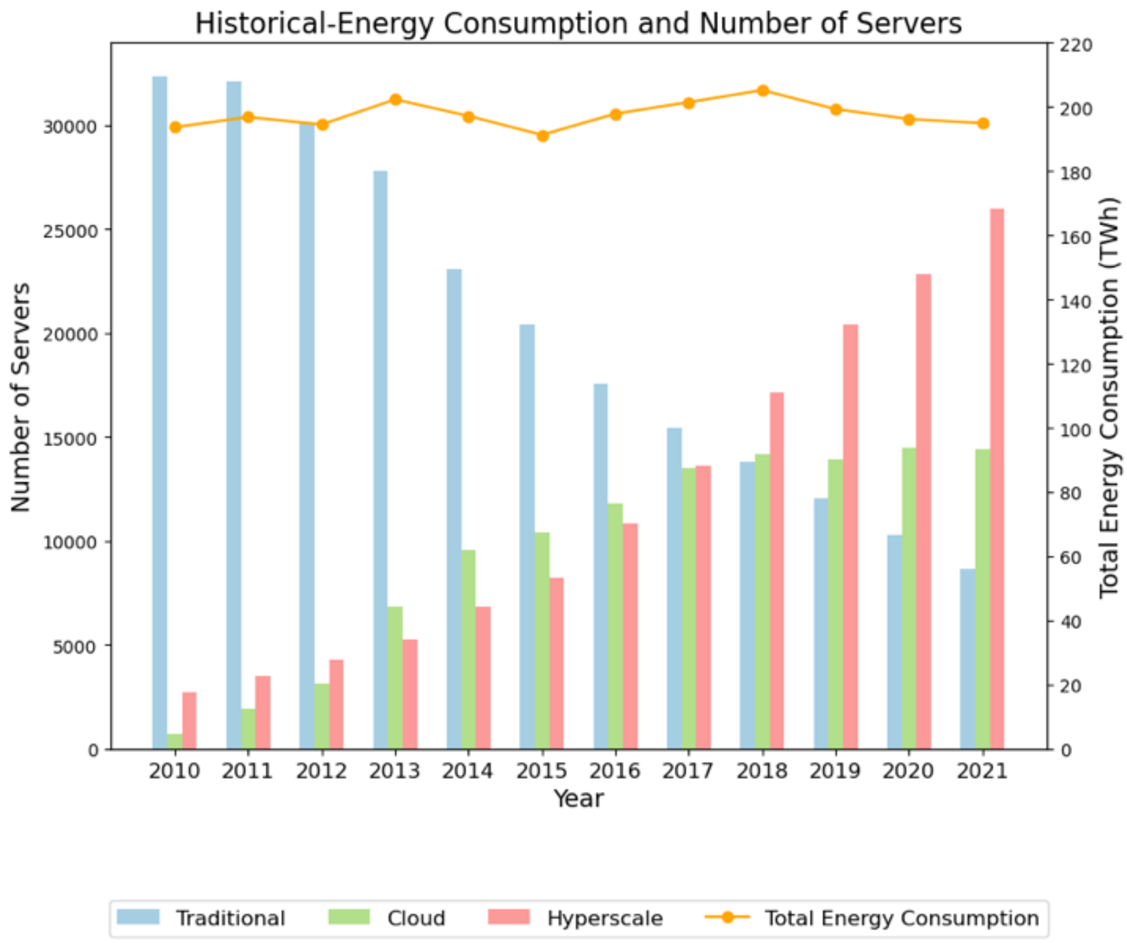

Fig 1. Historical energy consumption of data centres globally. Source: Cisco reports and Masanet et al.[6]

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

Since the launch of ChatGPT’s chatbot in November 2022, a highly concentrated market for genAI models has rapidly developed. In this initial discovery phase of its innovation lifecycle, genAI model performance improvements have been driven by a bigger-is-better logic, echoing the dominant scaling heuristic characteristic of other engineering industries.[7]

---

[6] Masanet, E., A. Shehabi, N. Lei, S. Smith and J. Koomey (2020). "Recalibrating global data center energy-use estimates." Science 367(6481): 984. 10.1126/science.aba3758

[7] Winter, S. (2008). "Scaling heuristics shape technology! Should economic theory take notice." Industrial and Corporate Change 17(3): 513-531.

The scale of genAI models coupled with widespread user uptake has in turn driven up data centre activity and energy needs, with some near-term projections indicating 3, 6, and even 8 fold increases in electricity consumption over the next 3-5 years. The International Energy Agency (IEA) projects global electricity consumption in data centres could exceed 1,500 TWh by 2030 if current trends persist. Older projections made before the genAI boom, like the one by National Grid the UK, show much slower increases [Fig 2].

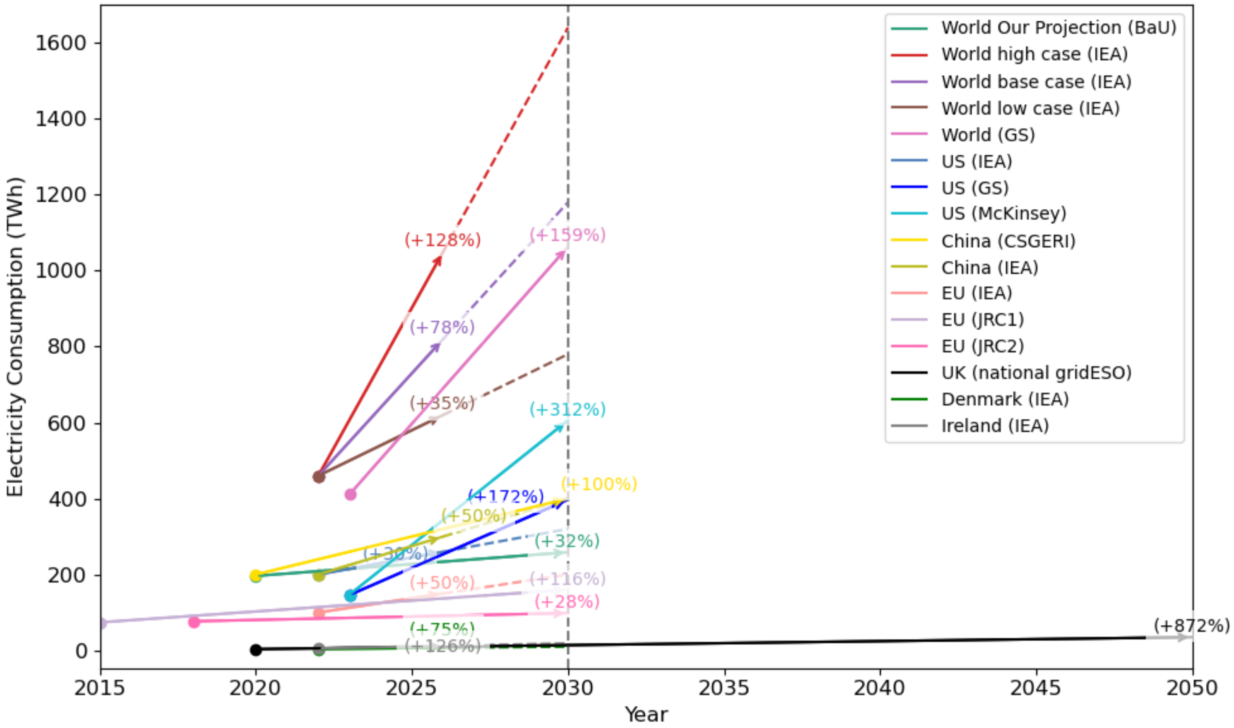

Fig 2. Projected increase in data centre electricity consumption in different regions. Circle markers show level in year projection was made; arrowheads show level in year projection ends (extrapolated linearly to 2030 with dotted lines). Source: Various.

For comparison purposes, we fitted a statistical model to data centre electricity consumption over the period 2010-2021. 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 to estimate future data centre electricity demand (labelled 'Our Projections' in Fig 2). Our projections are more conservative than others as they're calibrated to the historical period of data centre efficiency improvements and hyperscaling.[8]

Could a similar innovation response dampen near-term growth in data centre energy needs?

Current indications are yes. At the time of writing, launch of the Chinese chatbot, DeepSeek's R1, triggered a $400bn overnight loss in AI chip maker Nvidia's stock market valuation because DeepSeek outperforms other genAI models yet requires far fewer of Nvidia's high-performance GPU chips and so less electricity to run. Moreover, it's open source and is reported to have cost less than $10m to train (an order of magnitude lower than for other genAI models).[9]

In response to rising computational and energy costs, ChatGPT and other genAI models were already moving towards leaner code and smaller, task-specific models. The current generation of reasoning genAI models (including OpenAI's o1 and DeepSeek's R1) also spend more time on inference. How this shift in computational and energy needs from training large models to querying smaller models will affect genAI use is unclear[10] but it seems likely paying per use will become more common, helping to limit profligate demand.[iii]

Ultimately, the future energy needs of data centres will be determined by three main factors: (1) growth in demand for computation by AI and particularly genAI models; (2) innovation to develop more efficient AI models; (3) potential for continued efficiency gains in data centre hardware, software, and balance of plant. All three factors are uncertain. The availability of low-carbon power in areas data centres are located is a fourth factor that will determine how these future energy needs translate into GHG emissions.

First, will compelling AI use cases become widely adopted?

---

[8] https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about

[9] https://www.nbcnews.com/data-graphics/deepseek-ai-comparison-openai-chatgpt-google-gemini-meta-llama-rcna189568

[10] https://www.economist.com/business/2025/01/20/openais-latest-model-will-change-the-economics-of-software

GenAI holds significant promise - an early McKinsey estimate put this at \$2.6-4.4 trillion globally.[11] AI chatbots have become widely used both by service providers and consumers[12], driving up data centre activity. A recent genAI summit[13] in Oxford showcased commercial use cases dominated by marketing and sales to support and improve 'the consumer journey and experience'. Beyond these examples, the extent to which genAI will become embedded across social and economic activities is yet to be determined.[14] This includes the vision for agentic AI[15] in which task-specific AI assistants or 'agents' understand, represent and negotiate all parties' interests in online interactions with the capacity to execute resulting decisions on behalf of consumers or firms. Early stage markets for new technologies are often associated with hype. It seems likely that a 'shakeout' of AI use cases will occur over the medium-term, leaving those with compelling benefits that justify the computational and energy costs. Despite these uncertainties, near-term demand for data centre activity seems set to continue its phenomenal recent growth.

Second, will AI models become more energy efficient? And third, will data centres become more energy efficient?

McKinsey estimates that energy accounts for about 20% of the cost base for data centre business models.[16] So while energy efficiency is not the main cost driver, it is salient. As noted, current emphasis in the leading genAI firms is on optimising training and inference compute efficiency rather than (solely) improving performance through scaling. DeepSeek is an early example.

---

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

[12] https://www.demandsage.com/chatgpt-statistics/

[13] https://www.oxgensummit.org

[14] McKinsey (2024). The state of AI in early 2024: Gen AI adoption spikes and starts to generate value. San Francisco, CA, McKinsey & Company. https://www.mckinsey.com/~/media/mckinsey/business%20functions/quantumblack/our%20insights/the%20state%20of%20ai/2024/the-state-of-ai-in-early-2024-final.pdf?shouldIndex=false

[15] https://hbr.org/2024/12/what-is-agentic-ai-and-how-will-it-change-work

[16] McKinsey (2024). How data centres and the energy sector can sate AI's hunger for power. San Francisco, CA, McKinsey & Company.

It's likely that data centres will also continue to improve in energy efficiency. Koomey's Law is the energy efficient analogue of the better known Moore's Law that charts exponential improvements in computational power observed in microprocessors.[17] Koomey's Law charts the exponential improvements in the energy required per unit of computation. It helps explain the flat energy consumption trends in data centres historically [Fig 1a]. However, it was observed originally from data from the 1950s up to the early 2010s. Whether exponential improvements in the energy efficiency of computation[18] can continue through new AI chip architectures and less software bloat is an area of debate and innovation but signs are positive.[19][20] 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 drive down PUEs by improving cooling efficiency: best-in-class data centres have PUEs <1.4 and in some cases as low as 1.1 by making use of 'free' cooling from natural heat sinks like lakes, snowpacks, or the ground. The industry-led Climate Neutral Data Centre Pact targets PUEs of 1.3-1.4 (lower in cooler climates) for all new data centres.[21] However it remains uncertain how widely diffused these examples of best practice will become, particularly as data centre design and construction decentralises out of the market leading firms and regions.

Fourth, will sufficient and additional low-carbon power be available to meet data centre needs?

Tech companies have strong GHG emission reduction commitments, deep pockets, and are relatively price insensitive when it comes to electricity. These three criteria could spur a nuclear renaissance. Microsoft have signed a 20 year power purchase agreement from a nuclear reactor that was mothballed in 2019 for being uneconomic. (The other reactor at the same Three Mile Island nuclear plant has been shut down since its 1979 leak). Meanwhile Google and Oracle have reportedly ordered a series of

---

[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. 10.1109/MAHC.2010.28

[18] https://www.iea.org/data-and-statistics/charts/efficiency-improvement-of-ai-related-computer-chips-2008-2023

[19] https://www.latitudemedia.com/news/catalyst-can-chip-efficiency-slow-ais-energy-demand/

[20] 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. 10.1007/s10586-024-04767-y

[21] https://www.climateneutraldatacentre.net

small modular nuclear reactors (SMRs) to power their data centres.[22] The UK government has also just announced a commitment to both data centres and SMRs.[23] However, there is a mismatch in timeframes between the <2-3 years for data centre permitting and construction and the much longer time need to permit and build new nuclear capacity including SMRs. In contrast, large-scale wind or solar plants can be built in <3-5 years, 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. The availability of both surplus renewables and available network capacity varies strongly by location. Currently there is some evidence that capacity factors of fossil fuel plants are being increased to serve data centre demands, in the absence of sufficient renewable power.[24]

This raises an important fifth question. Will new data centres be built near demand for computation or near available and accessible low-carbon power?

The answer depends on the relative strength of different economic incentives. Locating data centres near cities or industrial areas[25] reduces latency and avoids the need for new ICT network infrastructure. Locating data centres in areas with sufficient electricity network capacity avoids permitting and connection delays. Locating data centres in areas with freely available heat sinks (or in cooler climates) helps reduce PUEs and so data centre energy needs. Locating data centres in areas near to available or potential low-carbon electricity helps meet net-zero commitments, and avoids the need for new power transmission infrastructure which is costly.

There is evidence these latter energy-related incentives are changing location decisions. States in the US like Iowa, Wyoming, and Ohio are

---

[22] https://www.ft.com/content/29eaf03f-4970-40da-ae7c-c8b3283069da

[23] https://www.world-nuclear-news.org/articles/uk-government-considering-role-for-smrs-in-ai-expansion

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

[25] OIES (2024). Global Electricity Demand: what's driving growth and why it matters? Oxford Institute for Energy Studies.

hosting large new data centre investments as traditional hubs like Virginia, California, and Arizona experience delayed network connections.[26]

Policies also influence location decisions: relatively large numbers of data centre in Ireland are the result of a favourable tax regime. Regulatory changes to how benefits, risks and costs are allocated between data centres and other large loads should also be expected, particularly where low-carbon generation or network capacity is scarce.[27]

Taking all these factors into account, it seems highly probable 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 [Fig 2a]. This is also shown in our own projections, calibrated to historical trends [Fig 2b].

The GHG consequences are important but manageable - electricity can be decarbonised. The IEA projects 50% of global electricity will be low carbon by 2026, up from 39% in just a few years.

However, global averages mask important localised impacts. In Ireland and Denmark, the IEA projects data centres may account for 15-30% and 9% of total electricity consumption respectively.[28] Problems with grid capacity and congestion have led to de facto moratoria on new data centres in The Netherlands, Singapore, and elsewhere.

These localised implications of data centres for electricity networks will be shaped by competing forces that should help address near-term issues.

Current 'hype' about exorbitant energy needs [Fig 2a] tends to omit consideration of these dampening factors: network connection

---

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

[27] Riu, D. Smiley, S. Bessasparis, K. Patel, "Load Growth Is Here to Stay, but Are Data Centers?: Strategically Managing the Challenges and Opportunities of Load Growth," Energy and Environmental Economics, Inc., July 2024. Available at: https://www.ethree.com/

[28] IEA (2024). Electricity 2024: Analysis and forecast to 2026. Paris, France, International Energy Agency.

constraints, a regulatory response to allocate cost and risk between data centres and other ratepayers, and location decisions.[29]

Indirect energy impacts: AI applications across sectors.

One of the co-founders of Climate Change AI, David Rolnick, likens AI to a hammer: "*The primary impact of a hammer is what is being hammered, not what is in the hammer. Just as the tool can smash things to bits or pound in nails to build a house, artificial intelligence can hurt or help the environment*".[30]

As this quote suggests, the magnitude of AI's indirect impacts is considerably larger than its direct energy footprint. Indirect impacts are the net outcome of energy-saving efficiency, substitution, and optimisation effects offset by energy-increasing rebound, induced demand, and intensification effects.[iv]

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).[v]

Which side the double-edged sword falls will vary for each AI application and its use context.

Some AI applications are unambiguously beneficial for energy and GHG emissions.[vi] Balancing supply and demand in real-time on electricity networks with intermittent renewable generation and distributed storage and flexibility assets is increasingly complex and data-intensive: hallmarks of an AI use case. There are numerous examples of AI models being used to improve power system optimisation, scheduling and dispatch, in support of the low-carbon transition.[31]

---

[29] Riu, D. Smiley, S. Bessasparis, K. Patel, "Load Growth Is Here to Stay, but Are Data Centers? Strategically Managing the Challenges and Opportunities of Load Growth." Energy and Environmental Economics, Inc., July 2024. Available at: https://www.ethree.com/

[30] https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions/

[31] 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. 10.1038/s41558-022-01377-7

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.[32]

Most AI applications sit on neither of these unambiguous poles. Whether their use can potentially help reduce energy or GHG emissions depends on rebound effects and whether they can be managed.[vii]

For example, digital twins and smart control systems help optimise 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.[viii]

The same applies to AI and now genAI shaping consumer preferences on retail platforms and through influencers[33]. Clothing and fashion exemplify both sides of the double-edged sword:[ix] reuse platforms like Vinted and fast fashion retailers like Shein have both seen phenomenal recent market growth.[x]

Advocates of 'AI for Good' inevitably focus on the best case world of AI applications with limited or no induced demand or scale effects. Assessments in this vein estimate 5-15% energy saving benefits from digitalisation.[34]

Our own evidence synthesis on the indirect impacts of numerous AI-enabled applications in buildings, transport, and industrial sectors compared best case with worst case worlds. Fig 3 gives an example for transport-related applications. This makes clear the double-edged sword. Most applications have indirect impacts on energy that can be both

---

[32] 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. [doi: 10.1007/s13202-021-01302-2]

[33] https://www.forbes.com/sites/lesliekatz/2023/11/24/this-ai-generated-influencer-can-pull-in-10000-euros-a-month/

[34] GESI (2022). Digital with Purpose: Delivering a SMARTer2030. Brussels, Belgium, Global Enabling Sustainability Initiative (GESI). digitalwithpurpose.gesi.org

energy-saving [blue dots] and energy-increasing [red dots]. 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 optimised vehicle routing, driving, and occupancy are offset by increases in vehicle trips and distances.[35]

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.

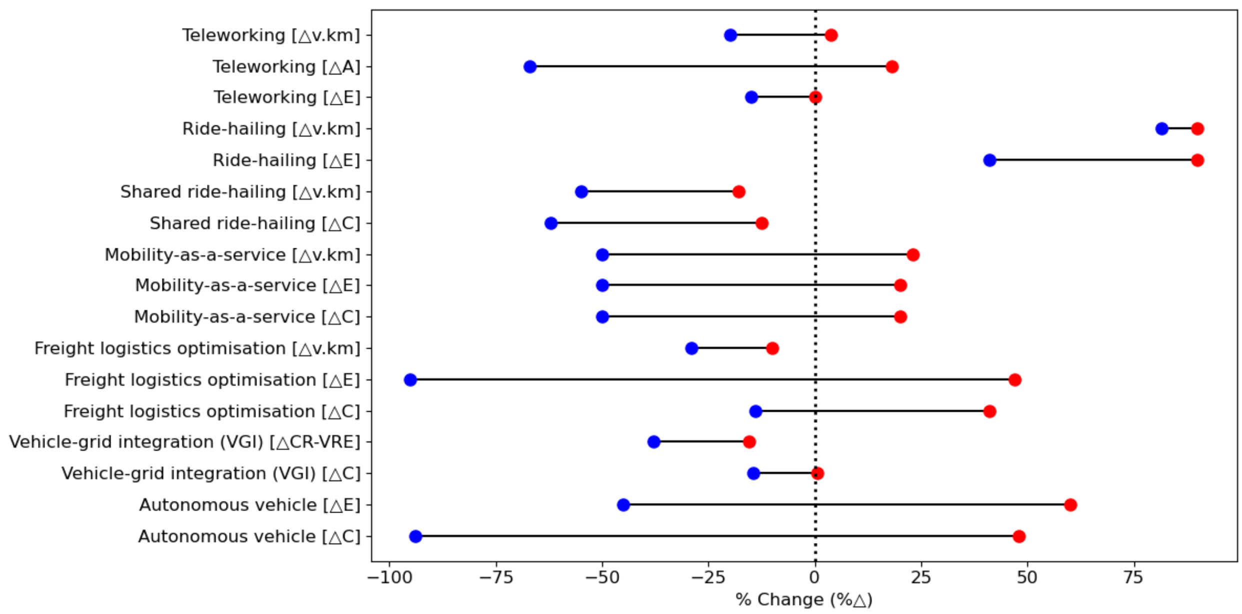

Fig 3. Indirect impacts of transport-related digital applications activity, energy, or emission outcomes. ΔA or Δv.km = change in activity or vehicle.kilometres travelled; ΔE = change in energy use; ΔC = change in carbon emissions; ΔCR-VRE = change in curtailment rate of variable renewable energy. Source: Wilson, C. et al. (2024). Evidence Synthesis of Indirect Impacts of Digitalisation on Energy and Emissions. 2024 10th International Conference on ICT for Sustainability (ICT4S). doi.org/10.1109/ICT4S64576.2024.00021

---

[35] 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. [doi: https://doi.org/10.1016/j.tra.2015.12.001]

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 are detrimental or beneficial, and are spread throughout application domains.

Policy considerations for direct and indirect impacts are therefore very different.[xi]

Direct impacts can be addressed through measures such as voluntary or mandatory energy performance standards for data centres[36], Scope 2 (electricity-related) emissions reporting requirements for data centre operators and tech companies, public sector procurement policies, land use planning and permitting, electricity network regulations governing connections for large new loads. Examples of all these policies have been implemented in the past few years in all major data centre locations including US, China, EU, Singapore. Recent IEA modelling projects only modest increases in data centre energy demand under combined policy assumptions.[37]

Indirect impacts can be addressed through energy or climate policies in the many different 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 or both. As Fig 3 shows, these outcomes have markedly different implications for energy.

Managing the risk that AI applications drive up energy demand or GHG emissions through changes in user behaviour are not part of current thinking on AI and energy governance. Environmental risks are recognised

---

[36] https://www.climateneutraldatacentre.net

[37] https://www.iea-4e.org/wp-content/uploads/2024/02/Policy-development-on-energy-efficiency-of-data-centres-draft-final-report-v1.05.pdf

in AI risk taxonomies[38][xii], but the energy and GHG-related dimensions are limited to direct impacts. AI regulatory frameworks like the EU's AI Act which is designed to preventatively assess and mitigate systemic risks also do not include the indirect impacts of AI applications on energy.[xiii]

Although many high-level policy narratives emphasise strong alignment between digital transformation and decarbonisation,[39] it is left to transport, buildings, industrial, and urban policies as instruments of climate governance to steer AI applications towards decarbonisation goals.[xiv]

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

A 2020 Royal Society report mooted the idea of an 'energy proportionality' test for new AI applications.[40] Those providing weak societal value but requiring large amounts of energy should be subject to regulatory guidance.

Sasha Luccioni from open-source AI company, Hugging Face, has similarly argued for selectively incentivising AI use cases that contribute towards the UN's sustainable development goals over ones that generate personalized ads for social media. UNESCO has recently published Recommendations on the Ethics of AI[41] that include the preference for data- and energy-efficient methods in order to mitigate both direct and indirect environmental impacts. The recommendations state that: "*where there are disproportionate negative impacts on the environment, AI should not be used*."

There are few signs yet of this happening, nor regulatory appetite to make it happen.[xvi] However, a necessary first step would be towards a more systematic tracking and reporting framework for the indirect impacts of AI

---

[38] MIT's AI Risk Taxonomy: arXiv preprint arXiv:2408.12622.

[39] https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/twin-green-digital-transition-how-sustainable-digital-technologies-could-enable-carbon-neutral-eu-2022-06-29_en

[40] Page 86: Royal Society (2020). Digital technology and the planet: Harnessing computing to achieve net zero. London, UK, The Royal Society. https://royalsociety.org/news-resources/projects/digital-technology-and-the-planet/

[41] https://unesdoc.unesco.org/ark:/48223/pf0000380455

on energy. The International Telecommunications Union has recently published standards for indirect emissions accounting which supports moves in this direction. More systematic reporting in turn enables investors like the international development banks to orient their green investment portfolios towards net GHG-reducing AI applications, and policymakers to differentially incentivise AI applications in domains where their net GHG impact is aligned with decarbonisation goals.[42]

Conclusion – Summary of key arguments and future outlook.

AI and now genAI applications are means not ends. Their use 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.[xvii]

Understanding AI and energy means "*grappling with both direct and indirect effects. Otherwise, the industry risks pinning its hopes on technical efficiency gains alone without recognizing the social, cultural, and economic contexts that materially shape technology uses*".[43][xviii]

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 so policy responses will be localised.

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

---

[42] ITU (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, Switzerland, International Telecommunication Union. https://www.itu.int/ITU-T/recommendations/rec.aspx?id=15030&lang=en

[43] Luccioni, A. S., E. Strubell and K. Crawford (2025). "From Efficiency Gains to Rebound Effects: The Problem of Jevons' Paradox in AI's Polarized Environmental Debate." arXiv preprint arXiv:2501.16548.

Endnotes.

---

[i] The proportion of total energy use as embodied energy in manufacturing is around one quarter overall, but up to one half for some end-use devices that have become much more efficient to operate. The materials needed to build data centres include small amounts of certain critical minerals like platinum and palladium, but these are negligible relative to global material flows for the clean energy transition.

[ii] The market intelligence firm, Gartner - associated with the 'hype cycle' concept - noted in 2008 that cloud computing could reconfigure how ICT services are provided and consumed In 2010, Microsoft launched its cloud computing platform, Azure, eight years after the launch of Amazon Web Services, the early runner in cloud computing among the big tech companies. Cloud computing applications and infrastructure have proliferated since (and continue to do so). [https://en.wikipedia.org/wiki/Cloud_computing].

[iii] GenAI models like DeepSeek that shift computational and energy needs from training to inference are challenging familiar operating models of tech companies based on high upfront costs and barriers to entry, but then close-to-zero marginal costs (monetised through advertising or economies of scope from positions of market dominance).

[iv] On aggregate, the net effect of digitalisation historically has been to increase energy consumption as absolute growth more than offsets relative productivity gains. [Lange, S., J. Pohl and T. Santarius (2020). "Digitalization and energy consumption. Does ICT reduce energy demand?" Ecological Economics 176: 106760. https://doi.org/10.1016/j.ecolecon.2020.106760]

[v] AI use cases that succeed in proliferating and becoming pervasively adopted will drive continued increases in demand for computation. This is one of the factors determining data centres' direct energy footprint. This more recent effect of AI is set against a backdrop of ever-deepening digital transformation and expanding digital infrastructure seen in almost all regions worldwide over the past 20+ years. [United Nations (2024), E-government Development Index: https://publicadministration.un.org/egovkb/en-us/About/Overview/-E-Government-Development-Index]

[vi] Climate Change AI, a global non-profit, comprehensively documents hundreds of other beneficial examples in industrial sectors, electricity systems, transportation, buildings and cities, agriculture, as well as in climate science and scientific research more generally. [https://www.climatechange.ai/summaries]

[vii] Most genAI use cases follow this same reasoning. Some will be unambiguously beneficial for saving or managing energy, or for reducing emissions. Most won't be. Whether deployed in scientific research, in consumers' online experiences, or in agent-to-agent interactions, genAI's contribution to decarbonisation efforts will depend on the specific ways and contexts in which it's used.

[viii] For example, AI can help accelerate scientific discovery of new materials. This may unearth new catalysts to improve the efficiency of electric vehicle batteries or carbon capture units. But it may also generate breakthroughs in the efficiency of fossil fuel extraction or synfuel combustion, improving the economic competitiveness of carbon-intensive energy resources.

[ix] It is possible that these can help reduce wasteful consumer choices or even steer consumers towards sharing economies or other types of more sustainable behaviour. But it is also possible that the ease and availability of online consumption matched, ratcheted up by expectations of immediate fulfilment and ever-shorter delivery times, is increasing unsustainable consumption.

[x] Already in 2013, over a third of what consumers bought on Amazon was based on product recommendation algorithms. Advertising is the main source of revenue for Google, Meta and other major AI players. [McKinsey (2013). How retailers can keep up with consumers. https://www.mckinsey.com/ch/~/media/McKinsey/Industries/Retail/Our%20Insights/How%20retailers%20can%20keep%20up%20with%20consumers/How_retailers_can_keep_up_with_consumers_V2.ashx]

[xi] Whereas the direct energy needs of AI computation is always detrimental, the larger indirect impacts of AI applications on energy needs can be detrimental or beneficial. These AI applications are diffuse. They span manufacturing, consuming, travelling, organising, recreating, doing science, farming, and numerous other activities.

[xii] MIT have compiled an AI risk repository, identifying 777 risks extracted from 43 risk taxonomies. Their domain taxonomy distinguishes socioeconomic and environmental risks as a domain. This is further subdivided into six subdomains: centralisation of power; inequality; economic and cultural devaluation of

---

human effort; competitive dynamics; governance failure (leads to ineffective governance of AI risks); environmental harm (e.g. energy and material footprint of AI hardware). This latter category of risk constitutes 2% of all risks identified, and does not include the amplification of energy-intensive behavioural patterns among firms or households as a result of using AI applications. [Slattery, P., A. K. Saeri, E. A. Grundy, J. Graham, M. Noetel, R. Uuk, J. Dao, S. Pour, S. Casper and N. Thompson (2024). "The AI risk repository: A comprehensive meta-review, database, and taxonomy of risks from artificial intelligence." arXiv preprint arXiv:2408.12622.]

[xiii] Digital governance including policies specific to AI like the EU's AI Act are designed to preventatively assess and mitigate systemic risks for consumer sovereignty, child safety, fair competition and taxation, and healthy democratic institutions. These systemic risks do not include the indirect impacts of AI applications on energy use and GHG emissions. These indirect impacts are accounted for in their respective application sectors.

[xiv] Unlike climate governance, AI governance currently involves very few actors (firms and states). But both climate and AI governance similarly depend on robust cooperative institutions, brokering, political agency, and the integration of justice and equity considerations to address important distributional issues and vulnerabilities.

[xv] A 'narrow' view of AI governance for climate mitigation argues climate governance should establish market incentives directing innovation activity leaving digital governance to 'just' tackle issues unique to the sector including misinformation, surveillance, abuse of market power, inappropriate use of digital media, labour market disruption, and the digital divide reinforcing other socioeconomic inequalities. In contrast, a 'broad' view of AI governance for climate mitigation argues that the direct and indirect impacts of AI on energy use and so GHG emissions also constitute a systemic risk that needs tackling at source: not just the direct energy footprint of ICT infrastructure, but also the indirect impacts of digital applications on energy-using activity at both micro and systems levels: from individual consumption choices on e-retail platforms up to labour productivity gains throughout industry and the economy.

[xvi] The continuing growth in cryptocurrency is an example - regulatory responses concern financial market stability not a GHG emissions footprint the size of Morocco. [Cambridge Bitcoin Electricity Consumption Index, https://ccaf.io/cbnsi/cbeci]

[xvii] Sasha Luccioni, climate lead at open-source AI company Hugging Face: "*The trajectory of AI's impact hinges on business incentives and market logics, governance and policymaking, and broader social and cultural norms …"* [Luccioni, A. S., E. Strubell and K. Crawford (2025). "From Efficiency Gains to Rebound Effects: The Problem of Jevons' Paradox in AI's Polarized Environmental Debate." arXiv preprint arXiv:2501.16548.]

[xviii] Digital governance's more fundamental and longer-term importance for decarbonisation efforts is to maintain and strengthen the social trust or 'glue' necessary for strong climate governance. Climate mitigation creates winners and losers. Dissensus and polarisation erode governments' capacity to implement robust energy transition policies. AI-accelerated information flows can amplify both the grievances of losers and the polarisation of political discourse. In so doing, AI may undermine progress towards net-zero with its clear long-term benefits but near-term transitional costs. This systemic risk of AI for energy use and GHG emissions is the most significant of them all.