

Governing AI Infrastructure on Decarbonizing Grids

Reda El Makroum^{*1}, Catherine Banet², Yee Van Fan³, Benedicte Fasmer Waaler⁴, Boris Gamazaychikov⁵, Sébastien Gros⁶, Ignacio Hernando Gil⁷, Magnus Korpås⁸, Aidan O’Sullivan⁹, Meysam Qadrdan¹⁰, Torbjørn Berg Strømstad¹¹, Sebastian Zwickl-Bernhard^{1,12}, Mark O’Malley¹³, and Hans Auer^{1,12}

¹Energy System Economics Lab, Institute of Net-Zero Energy Systems, Technische Universität Wien, Vienna, Austria

²Department of Energy and Resources Law, Faculty of Law, University of Oslo, Oslo, Norway

³Environmental Change Institute, University of Oxford, Oxford, UK

⁴Innovation Norway, Oslo, Norway

⁵Sustainable AI Group, Paris, France

⁶Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

⁷Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, Portugal

⁸Department of Electric Energy, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

⁹UCL Energy Institute, University College London, United Kingdom

¹⁰School of Engineering, Cardiff University, Cardiff, UK

¹¹Tydal Næringssselskap, Tydal, Norway

¹²Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

¹³Department of Electrical and Electronic Engineering, Imperial College London, London, UK

Abstract

AI-driven data center expansion is outpacing grid capacity, with developers increasingly bypassing the grid to build on-site gas plants. We propose a policy framework that ties grid access to flexibility provision, operational transparency, and tangible community and societal benefit, with Norway as the focal case.

Introduction

Since the release of ChatGPT in 2022, the race to scale generative AI has driven a sharp acceleration in global data center expansion. A new layer of demand, driven by large language model training and inference, is projected to double global data center electricity consumption by 2030 and triple it by 2035¹. Norway is one epicenter of this expansion. The country’s low electricity prices, cold climate, and nearly 100% renewable grid (see Figure 1) have attracted over €10 billion in committed hyperscale investment (facilities that exceed 100 MW of capacity).

^{*}Lead contact and corresponding author: reda.el.makroum@tuwien.ac.at

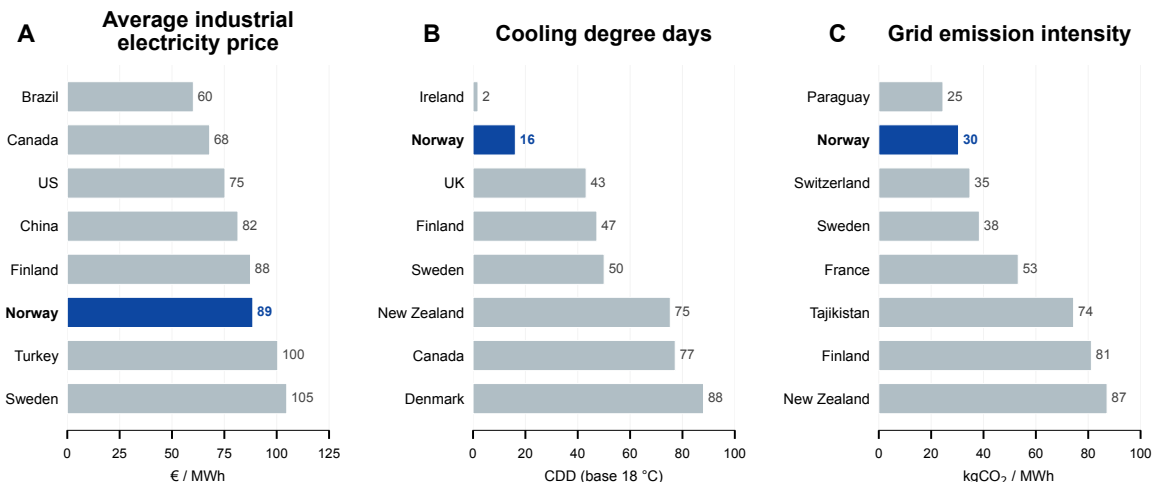


Figure 1: The main three factors driving data center investment toward Norway. Panels show (A) average industrial electricity price (2025), (B) cooling degree days, a measure of natural cooling potential (2020), and (C) grid emission intensity (2023). Lower values indicate more favorable conditions. Norway ranks near the top across all three among countries with sufficiently large grids to host at least one hyperscale data center. We measure grid size by annual electricity generation and set the threshold at 20 TWh. Below this scale, a hyperscale’s ~ 2 TWh annual demand² exceeds 10% of national generation and requires major reinforcement.

As of 2025, connection applications from data centers alone total over 7 GW, half of all new demand in the national grid queue³. The geographical distribution of in-queue projects in Norway reveals a clear priority. Projects concentrate in the southern regions despite these having the highest electricity prices (see Figure 2A). The reason is that these zones also concentrate the most transmission infrastructure, offering the fastest timelines to grid access. For data center developers, one factor outweighs all others: speed to connection.

The implications of this trade-off are playing out in the United States, which accounts for roughly 45% of global data center electricity consumption¹. Connecting to the grid now takes more than 5 years, pushing developers to build their own gas plants rather than wait. By the end of 2025, 39% of all natural gas power capacity under development in the U.S. was dedicated to data centers, up from 5% a year earlier⁴. [As of April 2026, further announced capacity totals 37 GW \(see Figure 2B\).](#) While most of it is unlikely to be built (as on-site gas is prohibitively expensive at scale), a partial buildout would still lock in fossil capacity for decades.

In one case, developers even bypassed environmental law. xAI deployed dozens of gas turbines on trailer trucks [in Southaven, Mississippi](#), to power its Colossus [data center across the state line in Memphis, Tennessee](#). These turbines, labeled as mobile units to circumvent permitting requirements, were operated without the required federal air permits and are now the subject of a lawsuit⁵ under the Clean Air Act. As long as connecting to the grid remains less viable than building around it, the pattern is likely to repeat.

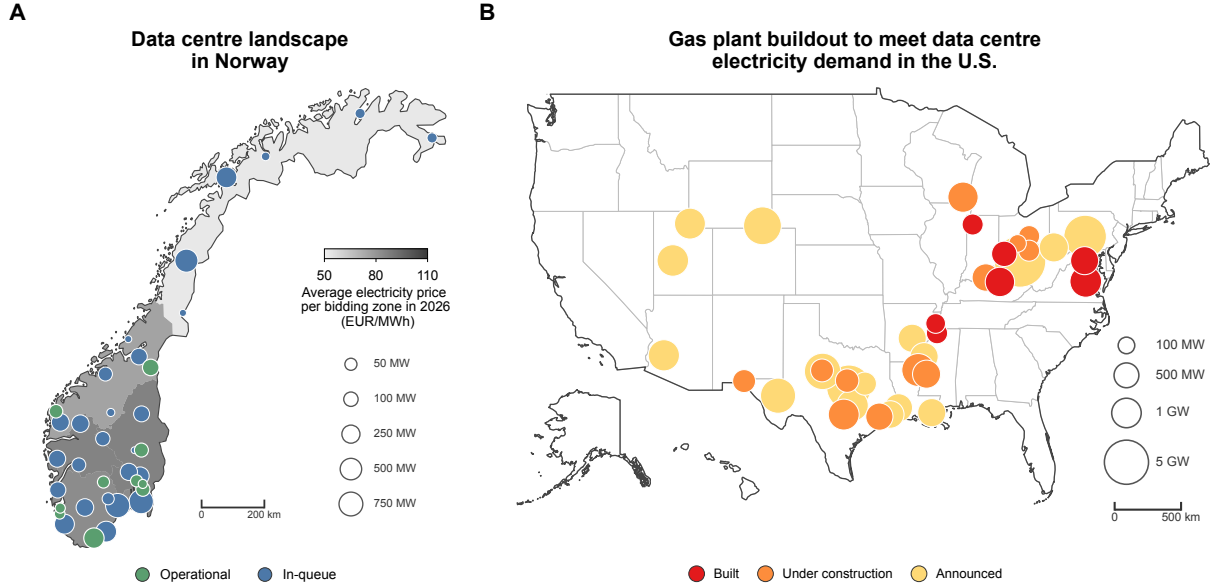


Figure 2: (A) Operational and in-queue data center capacity in Norway. Projects concentrate in the southern regions, which have the highest electricity prices but also the most transmission infrastructure, suggesting that faster grid access outweighs operating cost. (B) Built, under-construction, and announced gas-fired capacity serving data centers in the United States (including both grid-connected and on-site generation). About 12 GW is already operating or under construction, with a further 37 GW announced. The announced share is unlikely to fully materialize given the prohibitive cost of on-site gas at scale, yet a partial buildout would still lock in fossil capacity. Data as of April 2026.

Against this backdrop, we argue that a sustainable data center buildout requires tying grid access to flexibility provision, operational transparency, and tangible benefits for both host communities and society at large. The community dimension is particularly urgent: in 2019, public outcry forced Norway to halt all new onshore wind licensing for three years. As of 2026, the country is on track to miss its 2030 wind objective of 25 TWh, with only a single new wind farm built since the moratorium ended⁶. Without tangible local benefit, data centers risk becoming the next flashpoint.

Grid access as incentive for flexibility

Data center workloads are flexible: training can be shifted in time, while inference can be curtailed at peak demand⁷. This makes it possible for operators to provide flexibility, for instance by reducing consumption during peak hours (temporarily limiting GPU power) or by coordinating with virtual power plants (many small flexible resources pooled and run as a single plant). The technical potential exists, yet it remains not only untapped, but largely absent from data center planning. When applying for connection, developers prioritize uptime above all else. The result is agreements that lock in more capacity than facilities draw in reality (over 20% sits idle on average)².

Flexible connection agreements (FCAs) offer a way to close this gap. Under an FCA, a data center’s grid connection is split into firm capacity (guaranteed at all times) and curtailable capacity (which the operator can reduce during system stress). Evidence suggests that the disruption is modest. A modeling study of six U.S. data center sites found that transmission constraints required

curtailment for no more than 35 hours per year, and data centers operated with over 99% grid availability². How a data center meets curtailment is also worth specifying in the FCA: fallback should be restricted to clean resources [such as on-site batteries](#), rather than diesel or gas.

A recent amendment to the EU Electricity Directive (Article 6a)⁸ requires Member States to develop national frameworks for FCAs in capacity-constrained areas. While Norway is not yet bound under the EEA Agreement, Norwegian legislation already allows for conditional connection agreements⁹. The instrument, however, remains ad hoc and underused for lack of standardized terms.

Grid access reform should accompany the move toward FCAs. An accelerated connection track for data centers willing to accept curtailment and operate within existing grid constraints would reduce the dependence on transmission upgrades that take years to complete. This track would not replace the standard queue but run in parallel. The Southwest Power Pool in the U.S. has moved in this direction with its High Impact Large Load framework, which provides a conditional fast service track with interconnection agreements reached within 90 days¹⁰.

Reform must also address who gets priority. [Norway's](#) queue, which was originally managed on a first-come-first-served basis, shifted to maturity-based criteria in 2025 (requiring projects to prove they can finance construction to maintain their position). With the current investment landscape favoring data center buildout, competing battery, hydrogen, and ammonia projects lose their positions in the queue—projects that would tangibly advance the energy transition. Maturity criteria should thus consider the project's societal value alongside capital availability, so that the queue better reflects national decarbonization and energy security priorities.

[Once access is granted, ensuring flexible operation is the next challenge. This requires computing that can align with the grid's renewable generation. For Norway, that means aligning with hydropower, and specifically with the filling levels of more than 1,000 water reservoirs across the country. These levels swing with the seasons. They rise with the spring snowmelt and drop through the cold months when heating demand peaks. They also vary by region, and this is where the spatial flexibility of data center workloads matters. An operator with compute capacity where reservoirs are fuller can shift the workload there, relieving the local grid while keeping the computing running.](#)

Closing the data and transparency gap

Implementing FCAs requires data that grid operators do not currently have. The diversity of workloads, from sustained training runs to volatile inference, produces demand patterns that existing planning tools are not designed for. For this reason, data center load profiles remain poorly understood. Grid operators and energy researchers faced the same challenge with wind two decades ago, and addressed it through the IEA Wind Task 25, an international collaboration that pooled operational data and built shared methods for integrating variable generation. Data centers now require a similar effort.

No actor in the value chain has an incentive to disclose operational data. Providers who rent out server space do not know what tenants run on their hardware, and operators who run their own facilities classify load characteristics as commercially sensitive, with some tech companies actively lobbying to keep them confidential¹¹. The concerns understandably hold for raw data. However, approaches exist to derive anonymized load profiles and hardware specifications that are sufficient for grid planning. Such disclosure would enable grid operators to forecast demand, assess flexibility potential, and design connection agreements that reflect real load patterns.

The nature of the workloads themselves is another point of concern, with growing pressure to regulate what data centers process and prioritize societally beneficial computations over less critical

workloads. But deep monitoring of content at scale is technically infeasible. And even if it were feasible, the computational overhead and privacy implications make it both costly and invasive.

Nevertheless, transparency in data center operation is a research and policy effort worth pursuing¹², and workload type declaration is an essential first step. Norwegian legislation already requires cryptocurrency mining declaration¹³, and broader instruments are emerging at the EU level. The Artificial Intelligence Act (Regulation (EU) 2024/1689), for example, empowers the European Commission to refine energy reporting methodologies for AI models through delegated acts (secondary legislation that refines existing rules without reopening the full legislative process), a mechanism that could be extended to require workload type disclosure.

Shaping community and societal impact

Data center projects have an impact that can tip either way for the people living nearby. Data centers create jobs, but primarily during the construction phase as operational facilities require few permanent staff. Construction also brings broader economic activity to the region, yet the grid expansion required to serve data centers is likely to drive up electricity costs for all connected households and businesses. Beyond economics, waste heat is also a key consideration. [Electricity consumed for computing can deliver heat as a second service to district heating networks, thus improving the overall energy efficiency of the data center.](#) All three aspects require deliberate consideration if host communities are to benefit.

In Virginia, home to the world’s largest data center concentration, a state audit [warned that the public could effectively end up paying for the grid expansion data centers require](#)¹⁴. To mitigate this impact, community benefit agreements should address all three dimensions: local employment commitments beyond the construction phase, municipal revenue sharing to mitigate electricity cost increases, and waste heat co-design with district heating operators.

These agreements should be negotiated at the siting stage, not deferred until after construction. Tydal, home to one of Norway’s largest data centers, illustrates this gap, with waste heat from the facility remaining underused since it began operating in 2022. Closing the data gap matters here too: cooling load profiles must be shared alongside electrical load data so that planners can evaluate waste heat integration before siting decisions are made¹⁵. [A natural next step for these agreements is to mature into integrated flexibility products that simultaneously support power balance, heat supply, and computing stability.](#)

Discussion

Taken together, the instruments proposed converge into a governance framework (see Figure 3) [that cleanly integrates data centers into the grid, and for Norway, makes another wind experience less likely.](#)

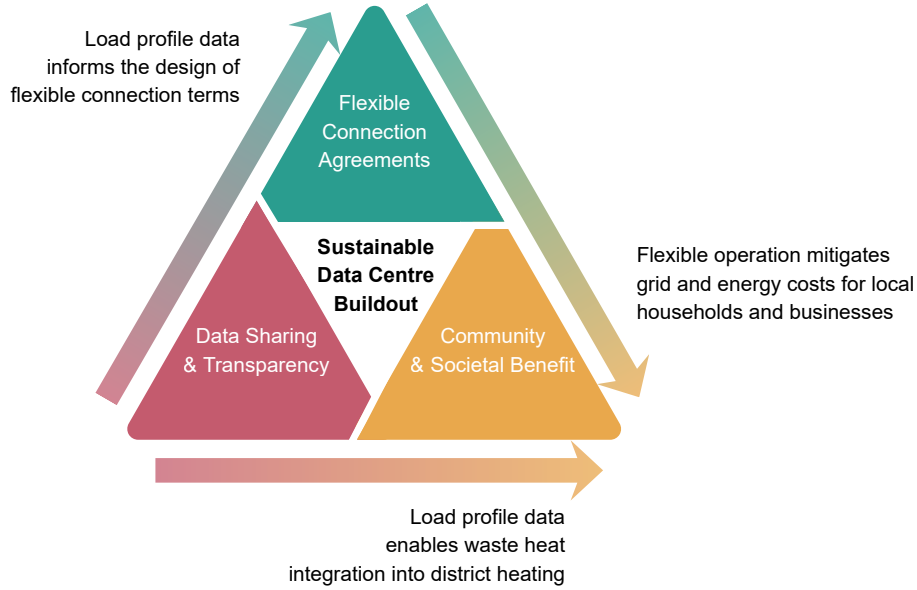


Figure 3: Proposed governance framework for sustainable data center buildout.

This framework is a starting point rather than a complete answer, as the governance challenge extends further. For individual AI models, the energy footprint must be reliably quantified, as model efficiency directly determines the infrastructure load that grids must absorb. At the sectoral level, the productivity gains from AI agents entering and transforming industries will drive economic growth, thus creating an indirect energy footprint that may prove equally significant. [A more hopeful dynamic is also at work. Guaranteed demand from data centers is helping de-risk clean technologies that have long struggled to attract financing. Geothermal systems and fuel cells are early beneficiaries that could mature faster as a result.](#)

The question of how long this investment wave lasts will ultimately be the most defining. Speculation about a potential AI bubble has made the parallel to the dot-com crash of the early 2000s hard to avoid. At the height of the 1990s internet boom, a massive fiber buildout deployed far more capacity than demand required at the time, with 95% initially unused. When the crash came, many of the companies that laid this fiber did not survive, and the vast majority of installed capacity sat unused for years. But as internet adoption grew through the 2000s, that same infrastructure became the backbone of the modern internet. So, whether AI demand accelerates or plateaus, the infrastructure being built will outlast the investment cycle that financed it, and the governance decisions made today will shape whether we future-proof the grid or derail the energy transition.

Acknowledgements

The authors alone are responsible for the views expressed in this publication, and such views do not necessarily represent the views, decisions, or policies of the institutions with which they are affiliated. This commentary draws on insights from a cross-disciplinary workshop held during the NTNU Energy Transition Week in March 2026. The authors thank the NTNU Energy Transition Initiative for its support. The authors also thank Antonia Golab and Lukas Kranzl for useful discussions and manuscript revision.

Author contributions

R.E.M. conceived this commentary and wrote the first draft. All authors contributed to discussions at the AI & Data Centres in the Energy Transition workshop and to manuscript revision. R.E.M., S.G., M.K., M.O., and H.A. also contributed to organizing the workshop.

Data availability

Industrial electricity prices are from Eurostat, supplemented with national sources for non-EU countries. Cooling degree days data are from the World Bank. Grid carbon intensity and electricity generation data are from the Ember Global Electricity Review via Our World in Data (2023). Data center queue data for Norway are from Statnett. Norwegian bidding zone prices are 2026 averages from Energy Charts and Nord Pool. Operational Norwegian data center locations were compiled by R.E.M. from public sources. U.S. gas plant data are compiled from the Global Energy Monitor and the U.S. Energy Information Administration, [supplemented by press reporting](#).

Declaration of interests

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, R.E.M used Claude to support manuscript preparation. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- 1 International Energy Agency (2025). Energy and AI. www.iea.org/reports/energy-and-ai.
- 2 Camus Energy, encoord, and Princeton ZERO Lab (2025). Flexible Data Centers: A Faster, More Affordable Path to Power. Tech. rep. Camus Energy. www.camus.energy/flexible-data-center-report.
- 3 Norwegian Government (2025). The Data Centre Industry – A Sustainable Industry of the Future for the Digital Norway. www.regjeringen.no/en/documents/the-data-centre-industry-a-sustainable-industry-of-the-future-for-the-digital-norway/id3112356/.
- 4 Elliott, R. F. and Stevens, H. (2026). Why Tech Giants Are Ditching the Power Grid. *The New York Times*. www.nytimes.com/interactive/2026/03/18/business/energy-environment/data-center-energy-gas-generators.html.
- 5 Earthjustice (2026). NAACP Sues xAI for Illegal Pollution from Data Center Power Plant. earthjustice.org/press/2026/xai-sued-for-illegal-power-plant.
- 6 Det Norske Veritas (DNV) (2024). Energy Transition Outlook – Norway 2024. www.dnv.com/publications/energy-transition-norway-2024/.

- 7 Oviedo, F., Kazhamiaka, F., Choukse, E., Kim, A., Luers, A., Nakagawa, M., Bianchini, R., and Lavista Ferres, J. M. (2026). Energy use of AI inference, efficiency pathways, and test-time scaling. *Joule*, p. 102430. <https://doi.org/10.1016/j.joule.2026.102430>.
- 8 European Parliament and Council of the European Union (2024). Directive (EU) 2024/1711 of the European Parliament and of the Council of 13 June 2024 amending Directives (EU) 2018/2001 and (EU) 2019/944 as regards improving the Union’s electricity market design. eur-lex.europa.eu/eli/dir/2024/1711/oj/eng.
- 9 Norwegian Energy Regulatory Authority (RME) (2024). Tilknytning med vilkår om utkobling. www.nve.no/reguleringsmyndigheten/regulering/nettvirksomhet/nettilknytning/dette-er-leveringsplikten/tilknytning-med-vilkaar-om-utkobling/.
- 10 Southwest Power Pool (2025). Southwest Power Pool board approves accelerated pathway for large load connection. spp.org/news-list/southwest-power-pool-board-approves-accelerated-pathway-for-large-load-connection/.
- 11 The Guardian (2026). Microsoft and other U.S. tech firms lobbied EU for secrecy rules on data centre emissions. www.theguardian.com/technology/2026/apr/17/microsoft-us-tech-firms-lobbied-eu-secrecy-rules-datacentre-emissions.
- 12 Hankendi, C., Coskun, A. K., and Sovacool, B. K. (2025). Why transparency matters for sustainable data centers and carbon-neutral artificial intelligence (AI). *iScience* 28.11, p. 113705. <https://doi.org/10.1016/j.isci.2025.113705>.
- 13 Norwegian Ministry of Digitalisation and Public Governance (2024). Lov om elektronisk kommunikasjon (ekomloven), Section 3-7 – Registreringsplikt for datasentertilbydere. lovdata.no/lov/2024-12-13-76.
- 14 Joint Legislative Audit and Review Commission (JLARC) (2024). Data Centers in Virginia. jlarc.virginia.gov/landing-2024-data-centers-in-virginia.asp.
- 15 Romanov, D., Chakraborty, I., and Holler, S. (2025). Comparative analysis of scenarios of data center waste heat utilization for district heating networks of different generations. *Energy Conversion and Management* 334, p. 119856. <https://doi.org/10.1016/j.enconman.2025.119856>.