Towards a Sustainable and Competitive AI Economy

Recommendations to the German Federal Environment Ministry

Prof. Dr. Udit Gupta[1], Prof. Dr. Philipp Hacker[2], Prof. Dr. Lynn H. Kaack[3,4], Prof. Dr. Emma Strubell[5], Prof. Dr. Aimee van Wynsberghe[6]

April, 2026

[1]Cornell Tech, New York City
[2]European New School of Digital Studies, European University Viadrina, Frankfurt (Oder)
[3]Hertie School, Berlin
[4]Climate Change AI, Portland
[5]Carnegie Mellon University, Pittsburgh
[6]University of Bonn, Bonn

Executive Summary

Artificial intelligence (AI) is increasingly transforming economic activity, public administration, and critical infrastructure across Germany and Europe. This transformation raises urgent questions not only about innovation and competitiveness but also about environmental responsibility and the governance mechanisms required to address it.

A central argument of this report is that **environmentally sustainable AI and economic competitiveness are not in tension – they are mutually reinforcing**. Germany's industrial AI landscape, driven by small and medium-sized enterprises (SMEs) and grounded in task-specific, specialized models, is already well positioned to benefit from this alignment. While significant use cases for large generative AI models exist, the most important AI applications in German industry – predictive maintenance, quality control, machinery control, process optimization – exemplify effective use of small, specialized models rather than large, computing-intensive generative systems. These specialized models consume less energy, require less hardware capacity, match the operational realities of Germany's industrial base, and provide digital sovereignty if locally developed and deployed.

At the same time, the computing-related environmental costs of AI, particularly those of very large generative models and reasoning-intensive workflows, are substantial and growing. Further, the deployment of AI for sustainability purposes is lagging behind its potential. We argue that Germany is well positioned to capitalize on sustainable AI and the use of smaller and more specialized models. If addressed proactively, sustainable AI can become an area in which Germany establishes global leadership, a demonstration that technological innovation, economic efficiency, and environmental stewardship advance together. This report, the result of a joint effort among experts from computer science, AI research, law, public policy, and AI ethics, presents concrete policy recommendations to realize that opportunity.

Policy Recommendations

The recommendations are organized around four pillars.

I. Beyond Large Language Models and General-Purpose AI

1. **Dedicated funding** should be provided for the development and deployment of **specialized AI models and use cases**, to strengthen the Germany and European SME-driven AI ecosystem and promote inherently energy-efficient solutions.
2. **Targeted investment in alternative AI paradigms should foster progress** beyond current generative models, such as application-driven machine learning, neuro-symbolic AI, joint-embedding predictive architectures, neuromorphic computing, causal AI, embodied AI & robotics, and probabilistic programming.
3. **Public-private partnerships** should pool industry-scale data within secure frameworks, for instance through federated learning, while policy should foster ecosystems of specialized, transparent AI service providers, and facilitate access to fit-for-purpose compute.

II. AI Model Transparency for Efficient Markets and User Choice

4. Environmental reporting frameworks should be expanded to **mandate standardized, publicly accessible, and independently audited disclosures** of AI's computing-related impacts across the AI model lifecycle, explicitly including inference. The critical need for data to inform effective resource management and environmental protection outweighs the potential competitive disadvantages of disclosing metrics indirectly related to model size.
5. Beyond computing-related impacts, **the AI Act should be updated to mandate reporting of AI's application-related impacts by deployers** under Article 26, supported by centralized use-case tracking and public-private partnerships. This will enable more accurate modeling of how AI's applications impact the environment, and data-driven development of actionable pathways to maximize the net positive impact.
6. Users should have the **right to use a green model** by mandating that AI platforms offer users the option to default to the most energy-efficient model that provides the required modality, with advanced reasoning capabilities turned off by default so as to avoid unnecessary energy consumption associated with over-provisioned compute.

III. Sustainable AI Infrastructure

7. Environmental **reporting requirements for data centers should be strengthened to mandate public and independently audited disclosures**, which also account for the specific demands of large-scale AI workloads including both Scope 2 and Scope 3 greenhouse gas emissions**.** Critically, reported metrics should reflect actual operational conditions validated over a full twelve-month reporting period, rather than design estimates or self-reported figures.

8. Public investment should **fund and support the development of energy-efficient, task-specialized compute infrastructure suited to Germany's industrial needs,** enabling more efficient resource provisioning, longer hardware lifetimes, and the adoption of emerging sustainable semiconductor technologies. Beyond the design of individual facilities, investment should also support the development of grid-aware data center operations, including techniques that enable data centers to respond dynamically to real-time grid conditions through intelligent workload scheduling, cooling modulation, and the activation of on-site generation or storage resources.

9. **Ensure that data centers are subject to carbon pricing mechanisms and renewable energy additionality obligations.** Carbon pricing is an effective, market-driven mechanism to improve energy efficiency and promote renewable energy use, and exemptions in the interest of fast data center expansion, such as through electricity price compensation schemes, risk locking in inefficient infrastructure for decades. Alongside carbon pricing, AI companies and data center operators should bear an obligation to develop additional renewable energy capacity to offset the extra demand generated by their operations, preserving access to existing renewable supply for other crucial sectors of the economy.

IV. Effects on the Energy Transition

10. Regulatory changes with a narrow focus and short time-horizon may lead to long-term problems and a failure to meet energy transition targets. Potential reforms to core energy regulations **should preserve due process and fair procedures, instead of reacting to short-term demand like expanding data center capacities.**

11. We recommend that the German government **commissions a study on how German firms benefit from data center construction and operation in Germany and worldwide**, what environmental effects relate to that, and how sustainable solutions can be supported. This study can then inform the discussion of how Germany can pioneer specialized sectoral AI products that are beneficial for the energy transition and other aspects of environmental sustainability.

12. Fostering data center buildout in Europe with high environmental standards can create a market for innovative products from European companies. If incentives are set right, data center investments can **support renewable energy and storage technology industries and foster the local economy**, such as with wind energy development.

Introduction

Artificial intelligence (AI) is rapidly becoming a central component of Germany's, and Europe's, digital transformation. As AI systems are increasingly integrated into economic activity, public administration, and critical infrastructure, their development and deployment raise urgent questions not only about innovation and competitiveness but also about environmental responsibility **and the governance mechanisms needed for this**. Recognizing these challenges and opportunities, we present this document as a set of policy recommendations for advancing **sustainable AI for a competitive economy in Germany and in Europe**.

This report is the result of a **joint effort among experts from computer science, artificial intelligence research, law, public policy, and AI ethics**, all of whom are actively engaged in the development, study, and governance of AI systems. Drawing on our diverse areas of expertise, we offer **recommendations to the Federal Environment Ministry** aimed at supporting Germany's leadership in shaping an environmentally responsible digital future. While our perspectives and disciplinary approaches differ, we share a common concern: Expansion of AI infrastructure and applications must be aligned with environmental stewardship and long-term sustainability goals, and environmentally beneficial uses of AI need to be incentivized.

We acknowledge that many of the environmental challenges discussed in this report are not unique to AI; rather, they arise more broadly from processes of digitalization. Nevertheless, AI presents a particular dynamic moment within this broader technological transformation. Due to its growing computational intensity, technological potential, and strategic economic importance, AI offers a **unique opportunity to align environmental responsibility with economic ambition**. If addressed proactively, sustainable AI could become an area in which Germany establishes global leadership, demonstrating that technological innovation and environmental stewardship can advance together.

In this document, we use the term **sustainable AI** to refer to two complementary dimensions (van Wynsberghe, 2021). The first is AI *for* environmental sustainability, which focuses on the use of artificial intelligence to address environmental challenges, for example, improving energy efficiency, optimizing resource management, or supporting climate research (Rolnick et al., 2022; Hintz et al., 2025). We note, however, that AI can also be used in ways that undermine environmental sustainability. The second conceptual dimension is the sustainability *of* AI, which concerns the environmental footprint of AI itself, including the energy consumption, material extraction, hardware production, and end-of-life disposal associated with AI infrastructure and applications. Both dimensions are essential. The recommendations presented here aim to encourage policies that support the development of AI applications for environmental benefit and address the environmental costs embedded within AI systems and infrastructures.

For assessing and regulating the environmental impact of AI holistically, we need to distinguish three fundamental categories: the computing-related impacts, immediate application impacts, and system-level impacts (Kaack et al., 2022). The computing-related impacts (also "direct" impacts) include the energy and water consumption of large-scale computation, the material and resource requirements of hardware production, and the environmental implications of digital infrastructure across its lifecycle. While the focus of most discussion is on AI's direct, computing-related impacts, we also emphasize the importance of AI's application-related impacts (also "indirect" impacts), which arise from the use cases to which the technology is actually applied. While AI can be used to support environmental sustainability, it also has myriad applications that lead to further environmental destruction, such as advancing exploration and use of fossil fuels. By focusing on these manifold aspects related to sustainability, we aim to address a critical and often underexamined dimension of AI governance.

Through the recommendations that follow, we seek to outline practical steps through which Germany can support innovation in AI while ensuring that its development aligns with environmental responsibility. In doing so, we argue that **sustainable AI is not a constraint on technological progress, but a strategic opportunity**, one that can position Germany as a leader in responsible AI development and as a model for integrating digital innovation with climate and sustainability goals.

We organize our suggestions around four pillars, ranging from (i) a discussion of AI model types beyond large language models to (ii) environmental transparency concerning AI models for efficient markets and user choice, (iii) sustainable AI infrastructure, and (iv) suggestions relating to the economy and the energy transition more generally.

I. Beyond Large Language Models and General-Purpose AI

Background: From a technical, economic, and environmental perspective, it is important to distinguish different AI model types and modes of operation. While much of the attention recently has been focused on very large, so-called "general-purpose" AI models (GPAI) built within the paradigm of generative AI, significant work is being done in different AI paradigms, too. Most notably, smaller, task-specific AI models (e.g., non-generative AI used to perform classification and regression tasks) fulfill important roles in many industry use cases, from image recognition for cancer research to predictive maintenance. Specific real-world challenges in turn can inspire algorithmic innovation in AI, which is known as application-driven innovation (Rolnick et al., 2024). Even within generative AI model families, smaller, more efficient models are increasingly capable of taking on advanced tasks with adequate performance (Belcak et al., 2025; Ding et al., 2024; Ngo-Ho, 2025; Yang et al., 2025). Critically, the utility of additional computation, whether it arises from the model size, implementation or usage paradigm, is dictated by the task that the model is required to complete. Many tasks do not require or even benefit from the general or reasoning capabilities of general-purpose models (Cao et al., 2026; Grinsztajn et al., 2022; Pecher et al., 2025; Joshi, 2026), even though the frontier is evolving. Further, the more narrowly one can define the target task for a model, the more efficient it can typically be made. Therefore, in order to maximize the economic benefits of AI while minimizing its resource requirements, it is paramount to specialize models or use modes to tasks accordingly.

The key challenge consists in Germany's achieving and securing AI leadership in a part of the AI space that is relevant for its own economy and also respects planetary boundaries. Building on task-specific, non-generative and small generative models and systems, we believe that this is entirely possible. In this sense, environmental sustainability and competitiveness can be mutually reinforcing.

a) Frontier AI research with smaller models and different paradigms

Challenge: Large generative AI models are necessary for some tasks, but pose serious environmental challenges, particularly with respect to energy and water consumption (Kaack et al., 2022; Luccioni, Jernite & Strubell, 2024; Li et al., 2025). Frontier capabilities, such as reasoning and agentic AI, are bound to increasing the energy consumption further. According to the AI Energy Score project supported by Hugging Face and Salesforce (Luccioni & Gamazaychikov, 2025), reasoning models consume on average 30 times more energy than non-reasoning models; for specific models, the factor ranges from 150 to 700. Another key development are AI agents, which execute longer and more complex tasks by issuing many requests to Large Language Models (LLMs), and could drive another large increase in AI's energy consumption (Krupp et al., 2025). At the same time, a 2025 AAAI survey of 475 leading AI researchers found that 76% consider it unlikely that scaling current LLM approaches will yield progress towards significantly more powerful models (cf. AAAI, 2025, p. 63). The upshot is that the next breakthrough in AI may emerge from alternative paradigms, not from larger generative models of the current architecture (Bhuyan et al., 2024; de Melo, Koyejo & Mandt, 2025; see also Sarker et al., 2022). Furthermore, agents may be increasingly developed and deployed built on customized models, and be specialized for each industry vertical.

There are many other important developments in basic machine learning research that are not in the area of general-purpose models and crucial for realizing economic benefits of AI (Rolnick et al., 2024), many of which are particularly important for sustainability applications (Donti et al., 2022). Examples include physics-informed machine learning (ML), causal ML, interpretable ML, and uncertainty quantification. German institutions can be competitive in such application-driven areas of ML innovation.

Germany is well positioned to benefit from such a strategic shift. SMEs, the backbone of the German AI provider landscape, typically build specialized, non-generative or smaller models, or adapt larger generative models for specific industrial tasks (see also Section I. b).

Policy recommendations:

1. **Funding for specialized AI models and systems: Dedicated funding** for the development and deployment of **specialized AI models and use cases**, to strengthen the Germany and European SME-driven AI ecosystem and promote inherently energy-efficient solutions.
2. **Targeted investment in alternative AI paradigms:** Funding research and applications beyond current generative models, such as application-driven ML, neuro-symbolic AI, joint-embedding predictive architectures, neuromorphic computing, causal AI, embodied AI & robotics, and probabilistic programming. The SPRIND new computing concepts challenge is a welcome step in this direction (Bundesagentur für Sprunginnovationen, n.d.).

b) Fostering specialized AI applications for economic and sustainability benefits

Challenge: While LLMs have a distinct role to play, too, the most important AI applications in German industry, such as predictive maintenance, quality control, machinery control, process optimization (Heimberger, Jäger & Maloča, 2024, 6; Horvat et al., 2025; Bitkom, 2025b), often center on specialized tasks and rely on traditional machine learning rather than on large generative systems (see, e.g., Andrianandrianina Johanesa et al., 2024, 6). Specialized AI models, and not generative AI, are also behind almost all climate applications (Joshi, 2026; Hintz et al., 2025). Such specialized models matter for the German economy, and are often precisely the ones that use less energy for computing and do not require large hardware (e.g., Graphics Processing Units (GPUs)) capacities. A key challenge is that the deployment of AI across many sectors, in particular in sustainability-relevant applications, is slow. Barriers are manifold and include uncertainties relating to benefits and costs, a lack of data and digital infrastructure, as well as institutional and legal hurdles (Clutton-Brock et al., 2021; Hintz et al., 2026).

Policy recommendations:

1. **Reducing business risks:** Business risks can be reduced, for example, by fostering an ecosystem of AI service providers specialized in applications that are relevant to key sectors, such as the climate and the environment, and that are transparent about costs and benefits. This could also lead to companies building their own capacities in specialized AI applications and products emerging that can be marketed at scale. Policy-makers can also ensure that the regulatory frameworks for deploying AI facilitate deployment and avoid legal uncertainty, e.g. concerning liability. This concerns not only AI regulation but also sector-specific regulations. Platforms to share experience and knowledge about the integration of AI in industrial systems should also be facilitated.
2. **Fostering digital and data infrastructure:** For AI to be deployed in SMEs and for public-interest applications, data collection and sharing in and among these entities needs to be fostered with open data standards and data portals (Clutton-Brock et al., 2021). Also, public-private partnerships can help to pool and use industry-scale data within secure and confidential frameworks to reconcile data access with trade secrecy and EU General Data Protection Regulation (GDPR) compliance. This would unlock data for specialized models with a smaller environmental footprint or even developed directly for environmentally beneficial tasks. Policies to improve digitalization of sectors and enterprises more broadly are needed. Access to appropriate compute should be facilitated, whereby in many industry- and climate-related cases compute is not generally needed in the type and amount necessary for training LLMs, as in hyperscale data centers (Bhushan et al., 2025; Luccioni, Jernite & Strubell, 2024), but rather compute capacities that fulfill cost, security, and other criteria appropriate for the use case (see also Section III, Challenge c).

II. AI Model Transparency for Efficient Markets and User Choice

Background: Preliminary estimates indicate that AI's energy consumption and corresponding environmental impact are both non-trivial and escalating. For example, electricity demand from data centers in Europe is projected to nearly triple by 2035 (ICIS, 2025). This growth is already straining physical infrastructure; in Frankfurt, Europe's largest data center hub, data centers consumed nearly 30% of the city's total electricity as of 2022 (Klimareferat der Stadt Frankfurt am Main, 2022). The yearly operational (energy use) greenhouse gas (GHG) footprint of data centers in Germany is estimated at 6.5 million tonnes carbon dioxide equivalents (CO2e) as of 2024, and data center electricity capacity is expected to increase by about 75% to 4800 MW by 2030 (BMWK, 2025). However, **there is a widening gap between the rapid expansion of AI and the rudimentary data available to track its environmental effects, establishing a critical need for improved reporting and transparency.**

The goal of transparency is threefold: first, it enables informed decisions by market participants on whether to use AI or other technologies, and what AI models and modes to use more specifically; second, disclosures inform users when choosing a model under energy and financial costs and constraints; third, transparency forms a factual basis for energy forecasting and effective governance.

Current environmental accounting for AI focuses almost exclusively on computing-related (direct) impacts, and even then, the available data and standards are insufficient. Computing-related impacts include operational factors like the electricity required to train and run AI models and the resulting greenhouse gas emissions when that power is drawn from fossil fuel-based electricity generation. The operational energy arising from a single query to a specific generative AI model remains undisclosed for all leading proprietary models. Published figures are insufficiently documented to be reliable, emphasizing the need for rigorous, measurement-based reporting to build public trust and establish effective baselines. Computing-related impacts also encompass embodied effects, such as the carbon emitted during the manufacturing and disposal of specialized hardware like GPUs (Kaack et al., 2022; see Section 3 for further discussion). Meanwhile, there is virtually no systematic accounting of AI's application-related (indirect) impacts, which are those that arise as a result of how AI models are put to use in practice, which can lead to either expansion or reduction of GHG emissions. Proponents frequently cite application-related benefits, such as AI's potential to optimize smart grids or improve energy efficiency, to justify the technology's rapid expansion. Yet, without concrete data to ground these claims and with much less data characterizing negative indirect impacts such as accelerating fossil fuel extraction, the true net climate effect of AI, and corresponding decision points, remain unknown.

The current legislative landscape is beginning to address these issues, but significant gaps remain. The EU AI Act introduces baseline transparency measures, for example requiring developers of General Purpose AI models to document the computational resources used for training and the model's known or estimated energy consumption. However, these environmental provisions are confined to GPAI and high-risk systems, and they focus heavily on static estimates of the training phase (Alder et al., 2025; Hacker, 2024). This leaves a significant regulatory blind spot regarding the ongoing energy and GHG impacts of e.g. model inference, smaller specialized models, and the downstream integration of these systems into broader commercial applications.

Our understanding of AI's environmental impacts thus still largely relies on a fragile basis of voluntary corporate reporting and academic estimates, lacking the necessary granularity, scope, contextualization, and accuracy to inform user choice and political decision-making.

Transparency is crucial for market decisions and user choice, too. The environmental impact of generative AI depends not only on how models are trained and operated, but also on how they are deployed to, and used by, end users. Providers make critical design choices: which model to route a query to, whether to activate reasoning capabilities, and whether to offer AI-generated outputs by default, and others. These provider decisions co-determine the energy footprint of billions of daily interactions. Yet, these choices are, at present, often entirely opaque to users. EU law already recognizes that structural power asymmetries between providers and users demand regulatory intervention to preserve meaningful choice in areas such as data protection and privacy: for example, the GDPR requires that consent to data processing be freely given, specific, and informed (Art. 7 GDPR), and the Digital Services Act (DSA) mandates that very large online platforms offer recommender systems not based on profiling (Art. 38 DSA). Both regimes place the primary obligation on the provider, not the user, to create the conditions under which genuine choice becomes possible.

No equivalent framework exists for the environmental dimension of AI deployment. The result is a regulatory gap: users can neither reliably opt out of AI-generated features that they did not request, nor can they always select among models on the basis of energy efficiency data – even though the energy cost of a single query can vary by orders of magnitude depending on the model and configuration used (Luccioni & Gamazaychikov, 2025). This gap is particularly consequential: demand-side signals, when enabled through transparent choice architectures, could complement supply-side efficiency mandates and create market incentives for providers to compete on environmental performance. Hence, the challenge and recommendations rest on a dual premise: **the primary regulatory burden must fall on providers, who control the relevant design parameters; but regulation must simultaneously empower users with real autonomy to choose whether and how to engage with generative AI**.

a) General reporting and standardization

Challenge: The EU AI Act introduces baseline transparency for GPAI models and high-risk AI systems, but its environmental provisions are insufficient in clarity, scope, and accountability. While training a model requires a large initial input of electricity, increasing evidence points to inference, the everyday querying and running of the model, dominating the lifecycle energy use footprint (Luccioni, Jernite & Strubell, 2024; Jin, Wei & Brooks, 2025). However, the AI Act relies on static estimates of training energy, and only indirect data relating to inference energy (number of floating-point operations), which is insufficient to estimate actual energy use (Dehghani et al. 2022; Fernandez et al., 2025). This leaves a significant blind spot regarding the ongoing energy and GHG impacts of model inference and downstream deployment. In addition, there is a need to set parameters for how energy is measured in order to standardize reporting and make it comparable. The Code of Practice (CoP) for GPAI models leaves the reporting framework insufficiently vague. Critically, there is no specific measurement methodology provided for quantifying training energy or inference compute, leaving it to the reporting entity to decide and provide a description of the methodology.

Finally, data sparsity prevents researchers, policymakers, and downstream deployers from fully understanding the climate impact of specific AI tools. Broad corporate environmental transparency mandates such as Corporate Sustainability Reporting Directive (CSRD) only require reporting of high-level statistics on resource consumption, aggregated to the extent that they lack utility in this context. While the EU AI Act asks GPAI providers to share some technical documentation with downstream integrators, emissions and energy reporting is limited to government entities to government entities, based on claims of commercial secrecy. Moreover, these disclosures are not available to researchers or the general public (Alder et al., 2025), and lack independent audits. Yet, it is not clear that the potential competitive disadvantage that could arise from disclosing data indirectly related to AI model size outweighs the need for that data to inform effective resource management and environmental protection. And without independent verification, environmental figures are susceptible to underestimation, making it difficult to establish a reliable baseline.

Policy recommendations:

1. **Standardize and expand environmental reporting methodology:** Define clear standards for how relevant values should be reported, in particular in the AI Act, as leaving this to the reporting entity will result in incomparable data. Queries or tasks should, in turn, be standardized or measured on specific benchmarks to enable comparisons. Concretely, expand standardized reporting in the AI Act to explicitly include energy metrics for model development and use (inference) in addition to training, for example by means of Commission Delegated Acts. Regulators should establish a dynamic mechanism where standard-setting bodies regularly revisit and update reporting requirements. This ensures metrics remain relevant to the rapid evolution of the field and the changing capabilities of AI systems, effectively tracking absolute GHG emissions and water consumption alongside basic energy efficiency.
2. **Require broader access to disclosure of impacts:** The AI Act's CoP does not mandate disclosure of energy or computational requirements to downstream providers, let alone to the public. They are reported only to the AI Office and National Competent Authorities upon request. Regulators should improve transparency for downstream providers, ensuring that companies integrating foundation models into their own software have the data needed to calculate and report their own Scope 3 emissions. This would also allow downstream providers to select more efficient models as appropriate for their use cases, with corresponding economic benefits. This documentation should also be made publicly available to inform users and research through regular reporting.
3. **Establish independent audits of GPAI environmental impact reporting:** While corporate environmental reporting largely relies on voluntary self-assessments, data centers currently are placed under rather strict oversight to ensure theoretical models match actual operational and embodied GHG impacts. For the environmental audit of individual GPAI models, however, no such audits exist. One could build upon the mandatory third-party audit frameworks already established under the AI Act's GPAI CoP with respect to GPAI model safety, and with respect to infrastructure energy use and environmental impacts in the German Energy Efficiency Act (EnEfG).
4. **Leverage government procurement to promote compliance and transparency:** Current public procurement guidelines frequently lack environmental criteria for AI software and cloud computing services, focusing primarily on cost and functionality. Building on greater environmental transparency for AI models and services, Green Public Procurement (GPP) policies at the EU and national levels should be updated to establish environmental criteria and transparency requirements for AI systems. Government agencies should prioritize

purchasing AI solutions and cloud services from vendors that demonstrate compliance with established energy efficiency standards, renewable energy procurement, and transparent GHG reporting.

b) Indirect impact estimation and reporting

Challenge: The application-related/indirect impacts of AI are rarely measured, or qualitatively described by researchers and practitioners. As a result, any available estimates of the potential of AI for reducing emissions, and concerning positive environmental effects more generally, are based on rough assumptions by modelers and lack empirical grounding [cite IEA, Stern]. In addition, the impacts of AI applications that are detrimental to the environment, e.g. in the oil and gas sector, and the effects of AI driving growth of resource and energy consumption at the systems level are even more uncertain. No robust estimates of these effects exist and such negative application-related effects are often ignored in the debate.

Depending on the type of application, quantitative estimates of emissions effects may be easier or harder to obtain, which is why in some cases one may want to resort to qualitative descriptions of potential effects. However, we still lack fundamental data about where AI is deployed and what the environmental impacts are across sectors. Available data are predominantly reported by consulting firms with undisclosed methodology and a potential optimistic bias (Kaack et al., 2022). There is also no mandate to report these application-related impacts in current regulation; for example, the AI Act only mandates reporting of computing-related impacts.

Policy recommendations:

1. **Transparency on indirect impacts in the AI Act:** There is no mandate to report application-related impacts in the AI Act, which only requires reporting of computing-related impacts. Following up on repeated calls for the inclusion of such impacts in the AI Act (Kaack et al., 2021; Hacker et al., 2025), we suggest to at the minimum require a qualitative description of environmental effects of deployed high-risk AI systems. This should include effects on GHG emissions and water, and potentially other impacts. Such provisions should be added to the obligations of deployers in Article 26.
2. **Fostering impact assessment with public-private partnerships:** Government agencies should fund research and pilot projects to support companies in assessing both the computing-related and application-related environmental effects of their AI usage, with an explicit focus on identifying actionable pathways to reduce those impacts. This effort could be modeled after, or integrated with, existing frameworks like the Corporate Digital Responsibility (CDR) Initiative supported by the Federal Ministry of Justice and Consumer Protection (BMJV). By collaborating directly with companies that are already motivated to act in a digitally responsible manner, policymakers can test environmental reporting frameworks in real-world settings and establish proven, practical methodologies that can later be scaled across the broader industry.
3. **Understanding and monitoring application-related indirect impacts at scale:** Establish a centralized database and a regular reporting framework to track AI use cases over time. Entities deploying public-facing models should be required to report their AI use in different

application categories (e.g. with aggregate percentages). Concurrently, research funding should be directed toward surveying industry practitioners and supporting studies that use this data to report on and model broader sector- or economy-wide environmental effects. Blueprints for such regular reports have begun to emerge for other impacts of AI, e.g. on labor (Massenkof et al. 2026). This combined approach of mandatory transparency and funded research will allow regulators to link specific AI deployments upstream to their resource intensity and downstream to their true societal and environmental impacts.

c) Enabling user choice for sustainable AI

Challenge: Users today lack meaningful choice at two levels (Ebert, Gamazaychikov et al., 2026). First, they often cannot decide whether to use generative AI at all: providers increasingly embed AI features into search engines, email clients, and productivity suites through nudging mechanisms and dark patterns, without a prominent or persistent opt-out option. Even where such mechanisms exist – as with the DSA's right to non-personalized news feeds (Art. 38 DSA) – platforms tend to implement them in ways that are cumbersome and non-persistent, rendering the right ineffective in practice. Second, even when users do engage with generative AI voluntarily, they often cannot choose which concrete model processes their query – small or large, base or reasoning-enabled (Wang et al., 2024; Yu et al., 2024) – even though the environmental costs vary enormously. Reasoning modes, for example, use significantly more compute for a specific task to calculate a certain outcome (Luccioni & Gamazaychikov, 2025). This dual deficit deprives users of agency at the very moment when the environmental stakes of AI deployment are rising sharply, and it undermines the conditions for meaningful choice under which markets could reward more efficient AI solutions. Ultimately, user choice could also lead to users opting for less efficient models than automated provider-side routing of prompts to models; hence, salient information about the environmental consequences of model choice is paramount.

Policy recommendations:

1. **Right to use digital technology without generative AI:** AI-powered technology, including search engines, browsers, and coding tools, should provide users with a prominent option to disable AI-generated tools and return to traditional product use. For internet search, this implies the display of search results without overviews written by generative AI. Such an opt-out mechanism must be easily accessible, session-persistent, and clearly communicated.
2. **Right to use a green model:** When users do choose to engage with generative AI, they should have the right to select an environmentally optimized model. Providers typically deploy multiple models within a single model family for each prompt; the right to a "green model" would require providers to offer users the option to default to the most energy-efficient model that still provides the required modality. The AI Energy Score framework, with its star-based rating system, offers an example of a classification that regulators could incorporate by reference. Unless empirical studies show significantly superior energy efficiency performance by AI-based routing compared to active user choice, platforms should default to "reasoning off" and require an active user toggle to activate reasoning capabilities, accompanied by a visual indicator of the increased energy cost.

III. Sustainable AI Infrastructure

Data centers are fundamental to building an ecosystem within Germany that effectively leverages AI capabilities. They serve three crucial roles in the AI ecosystem: data storage and processing, training, and inference. As the backbone of both general-purpose and specialized AI, data centers underpin every stage of AI development and deployment from training foundation models to processing the billions of inference requests that deployed models receive every day.

Despite their central role in the AI ecosystem, data centers' scale and rapid expansion raises serious environmental and governance concerns. Data centers are large, complex infrastructure systems whose environmental footprint extends well beyond operational energy consumption (Gupta et. al. 2021; Lei et al., 2025).

The impact of data centers depends on how power demand is actually designed and utilized, how cooling systems interact with local water resources, how renewable energy is sourced and matched to consumption in time and location, which refrigerant is used, and how compute capacity is measured across increasingly heterogeneous hardware that spans general-purpose servers to specialized AI accelerators such as GPUs. Addressing these impacts requires not only accurate measurement and reporting, but deliberate choices in how data center infrastructure is designed, built, and operated. This includes investment in energy-efficient and grid-compatible technologies, the adoption of intelligent approaches to managing electricity demand in response to real-time grid conditions, and ensuring that new infrastructure is scaled in ways that reflect Germany's specific industrial AI needs rather than defaulting to the largest and most resource-intensive solutions available. Data center infrastructure represents a decades-long commitment of capital, energy, water, and land; getting these foundational choices right is therefore crucial to long term economic competitiveness, ensuring that Germany's digital infrastructure remains viable, adaptable, and cost-effective as both AI demands and environmental pressures evolve.

a) Beyond efficiency metrics: Transparent accounting on the absolute scale of impacts

Challenge: Current reporting frameworks, including established standards such as EN 50600-4, provide a meaningful foundation for assessing data center environmental performance. However, the value of these standards depends entirely on how faithfully they are implemented. In practice, reported values frequently reflect design estimates rather than actual operational measurements, and reporting is not always carried out in full compliance with the relevant standards. Even for relatively well-documented metrics like total data center energy use and associated GHG emissions, current data is insufficient to understand the status quo at the global level and inform basic forecasting, leading to wildly divergent projections. For instance, global data center energy estimates for 2030 vary by a factor of nearly 40, ranging from 210 TWh to nearly 8000 TWh (Kamiya & Coroamă, 2025). The challenge, therefore, lies not with the metrics themselves but with ensuring that they are applied consistently, accurately, and in ways that reflect real-world operational conditions. Data center growth must therefore go hand-in-hand with more rigorous, standardized, and transparent sustainability practices. For example, without rigorous transparency requirements the actual demand for new infrastructure remains obscure. Such data showed that Dutch data centers operate at roughly one third of their nominal capacity, suggesting that much of the data center infrastructure is underutilized (Schulze, 2025). Transparency can not only to prevent harm to local communities and the environment, but also ensure that Germany's expanding digital infrastructure is built on a foundation that is compatible with its long-term climate and energy transition goals.

The key metrics used to assess data center efficiency today are Power Usage Effectiveness (PUE), and Water Usage Effectiveness (WUE). PUE measures how efficiently a facility converts incoming electricity into useful computation, while WUE tracks water use relative to energy load. Both PUE and WUE are ratio-based indicators that describe relative efficiency rather than absolute consumption. Data centers can steadily improve PUE scores while their total energy draw grows year over year, because efficiency gains are routinely reinvested into larger models, more frequent retraining, or expanded deployment (Onitiu et al., 2026; Lei et al., 2025). This dynamic, known as rebound effect or Jevons' Paradox, means that improvements in efficiency do not automatically translate into a smaller environmental footprint (Luccioni, Strubell & Crawford, 2025; Wu, Raghavendra & Gupta, 2022). As AI workloads scale, the gap between efficiency ratios and actual impact becomes increasingly significant. A further limitation of PUE is that it captures only the relative energy consumption of the facility-level overhead (e.g., cooling systems, lighting, and power distribution) and provides no visibility into the IT stack itself, which accounts for the majority of total electricity consumption in AI-intensive facilities. Without metrics that reach inside the IT stack, efficiency reporting gives a systematically incomplete picture of where energy is actually going, such as training versus inference capacities and different infrastructure resources (e.g., data, accelerated compute, general purpose compute, networking).

Another salient challenge in the German data center ecosystem is that data centers predominantly operate on a co-location model, in which real estate companies build and operate data center campuses and rent space and electricity capacity to enterprise customers. Operators and the electricity consumers are different entities, which can create misalignment in capacity planning. Enterprise customers may reserve more electricity capacity than they ultimately consume, and co-location operators must in turn request that reserved capacity from the grid regardless of whether it is actually used. If enterprise customers were required to engage in rigorous capacity planning (specifying the electricity they need for their computational output and reporting measured utilization rates) co-location operators could make reliable and accurate load requests to the grid, and ease pressure on grid infrastructure.

Policy Recommendations:

1. **Mandate public disclosure and strengthen audit requirements:** Environmental information for data centers should be made publicly accessible, at minimum in aggregated form, with independent auditor access where confidentiality concerns apply. Existing reporting frameworks, such as those established under the German Energy Efficiency Act and the European Energy Efficiency Directive, should be strengthened to strictly account for the specific demands of large-scale AI workloads, including verification of Scope 2 and Scope 3 greenhouse gas emissions through mandatory third-party audits, rather than self-reported figures. Critically, reported metrics should reflect actual operational conditions validated over a full twelve-month reporting period, in addition to design estimates**.** Emerging frameworks for independently validated data center metrics, such as those developed by the Swiss

Datacenter Efficiency Association (SDEA), offer concrete models for what rigorous, independently verified reporting can look like in practice. Germany should position this domestic framework as a blueprint for broader EU-level reporting requirements, while actively participating in emerging international efforts to standardize environmental assessment for digital infrastructure through organizations such as NIST, ITU, and IEEE (p7100 standards), and within the regulatory context of the EU AI Act. By taking a leading role in shaping these standards, Germany can contribute to internationally recognized approaches to environmental transparency in AI infrastructure rather than adapting to frameworks designed elsewhere.

2. **Require rigorous reporting standards for design versus actual operational performance:** Utilization rates can lead to drastic differences between the rated design power and the actual operational power consumption. For the installed information technology power demand reported under the European Energy Efficiency Directive, operators should be required to report the actual installed electrical design capacity of their IT infrastructure, in kilowatts or megawatts, rather than figures derived from annual energy consumption. The German Energy Efficiency Act (EnEfG) specifies such reporting already, alongside actual operational consumption figures. In co-location facilities, enterprise customers should additionally be required to report their actual electricity utilization relative to contracted capacity, enabling co-location operators to make accurate and reliable load requests to grid operators. Utilization rates should be reported at server level as well as data center facility level to provide granular visibility of resource and energy use.
3. **Require granular and dynamic power consumption reporting:** Annual energy aggregates are insufficient to assess how data centers interact with local grids in real time. Operators should be required to report power consumption at finer time intervals, enabling regulators and grid operators to understand peak demand, renewable energy utilization, and the dynamic relationship between AI workload growth and grid stress.
4. **Require disaggregated reporting of training and inference workloads:** Training and inference represent fundamentally different types of computational demand. Training is intensive but periodic, while inference is lower intensity but continuous and growing rapidly at scale. In addition to model level reporting guidelines detailed in Section 1, operators should be required to report top-down training and inference workloads separately, including their aggregate demands. This would give regulators a far clearer picture of how AI use cases drive energy demand, consistent with calls in recent policy literature, including by the IEA and in academic proposals for AI environmental disclosure frameworks. Facilities where operators do not have visibility into granular workload demands, such as co-located data centers, may be exempt from these reporting guidelines.

b) Quantifying life-cycle impacts of data centers

Challenge: Existing reporting frameworks focus narrowly on operational energy consumption and carbon emissions from electricity. Data centers' environmental impact is much broader encompassing water consumption and electronic waste, and many more, as well as embodied impacts in hardware and building supply chains. These dimensions are largely unmeasured and only partially regulated.

Beyond energy consumption and carbon emissions, data centers consume significant volumes of water for cooling, as modern GPUs and specialized AI hardware generate substantial heat and require continuous cooling to maintain performance. In regions where water resources are limited or subject to seasonal variation, large-scale withdrawals can place real pressure on local reservoirs and water supplies. The environmental consequences vary considerably depending on location, cooling technology, and the scale of deployment, making standardized and locationally sensitive reporting essential (Lei et. al. 2025, Li et. al. 2025) . Further upstream, the production of AI hardware depends on global supply chains involving the extraction of critical minerals, energy-intensive manufacturing, and long-distance transportation, none of which are captured by operational energy metrics alone (Gupta et. al. 2021, Wu, Raghavendra & Gupta, 2022). At the end of their lifecycle, rapidly evolving hardware generations mean that older chips and servers are retired quickly, contributing to growing volumes of electronic waste. Developing detailed metrics across all of these dimensions is challenging, as impacts depend on complex supply chains and local environmental conditions. However, many of the relevant data points, including water consumption, cooling requirements, and hardware life cycle cost and replacement cycles (e.g., NVIDIA 2025, Falk et. al. 2025, Schneider et. al. 2025), are already tracked internally by data center operators. Greater transparency and standardized reporting could therefore build on existing industry data, incorporating it into lifecycle assessment frameworks that evaluate environmental impact from raw material extraction through to disposal.

Policy Recommendations:

1. **Strengthen water reporting to capture consumptive use, peak demand, and local water stress:** Current standards such as EN 50600-4 provide a meaningful foundation but their value depends on faithful implementation. Reporting requirements, however, should reflect real-world operational conditions rather than design estimates (Li et. al. 2026). Reporting should explicitly define water consumption as water lost through evaporation or not returned to its original source, with facilities using natural water sources required to report intake and return volumes separately. Annualized metrics such as WUE are insufficient on their own, as many facilities draw large volumes of water only on the hottest days of the year (Li et. al. 2025). Reporting should therefore also include peak on-site water demand in million liters per day, and should incorporate water stress-adjusted metrics, which account for local water availability rather than treating equivalent volumes of consumption as equivalent in impact regardless of location.

2. **Improve reporting of energy and water consumption under the European Energy Efficiency Directive:** For water, reporting requirements should explicitly define consumption as water lost through evaporation or not returned to its original source, requiring facilities using natural water sources to report intake and return volumes separately, and to incorporate local and seasonal water scarcity context. For energy, Cooling Degree Days should be derived automatically from publicly available geospatial weather datasets based on registered facility location, rather than relying on operator self-reporting where a large share of submitted data is currently unusable, and temperature thresholds should be recalibrated for data center environments.

3. **Support development of AI hardware specific reporting standards:** Existing compute efficiency metrics, including the SPEC SERT 2 benchmark embedded in the EU Ecodesign

Regulation for servers (Lot 9), were designed for general-purpose servers and do not adequately capture the energy consumption and efficiency of GPUs and Application-Specific Integrated Circuits (ASICs). Germany should actively engage in the Lot 9 revision process to ensure that AI-specific hardware metrics are incorporated, while also supporting broader international efforts through NIST and IEEE to develop standardized measurement methodologies for AI hardware efficiency at the global level.

4. **Extend reporting to cover lifecycle carbon and e-waste, with independent validation:** Reporting requirements should expand beyond operational energy to include embodied carbon in hardware supply chains, hardware replacement cycles, and e-waste volumes. Environmental impact figures reported by data center operators and cloud providers alike should be subject to independent third-party validation rather than accepted as self-reported figures, as the current lack of verification makes it difficult for regulators, downstream users, and the public to rely on these figures for decision-making and emissions reporting.

c) Scaling AI compute while respecting global and local sustainability constraints

Challenge: Germany must continue investing in data center infrastructure to remain competitive in AI, but current regulatory frameworks do not ensure that this expansion aligns with national climate targets, local resource constraints, or the supply chain sustainability goals on which long-term infrastructure resilience depends.

The objective is to ensure that data centers are designed, built, and operated in ways that are compatible with Germany's climate targets, local resource constraints, and long-term energy transition goals. Germany faces a genuine tension here: evolving data center infrastructure is essential for maintaining competitiveness in AI, yet unconstrained infrastructure growth risks placing unsustainable pressure on electricity grids, water systems, renewable energy supply, and the global supply chains on which hardware production depends, with significant implications for embodied carbon and resource extraction.

Meeting this challenge requires not only better measurement but active policy support for sustainable infrastructure design and deployment. This means creating conditions in which energy-efficient and grid-compatible data center technologies are developed and adopted, waste heat is reused rather than discharged, renewable energy obligations are meaningful and enforceable, and new infrastructure is sited and scaled in coordination with local energy system planning. Improved transparency and capacity planning in the co-location market, as discussed in Challenge a), may also reveal significant unused capacity within existing infrastructure that is currently obscured by the gap between reserved and actual electricity consumption, potentially reducing the need for new construction in the near term. Germany's industrial AI needs, as discussed in Section I, also suggest that the country does not necessarily require the largest and most resource-intensive compute infrastructure, and that investment in specialized, efficient alternatives can simultaneously serve economic and environmental goals.

Policy Recommendations:

1. **Enforce performance-based efficiency standards over design-based approximations:** Going beyond more rigorous reporting standards for design versus actual operational performance (see also Section III b), efficiency requirements for data centers should be assessed against actual operational performance, not theoretical design scenarios. In practice, a significant gap can exist between reserved and actual capacity; for example, research in the Netherlands found that data centers operate at roughly one third of their total electrical capacity on average, suggesting that much of the infrastructure being built and connected to the grid is underutilized (Schulze, 2025). Closing this gap through performance-based standards would reduce unnecessary grid reservations, lower the pressure to build new facilities, and create guidelines for operators to use existing infrastructure more efficiently.

2. **Demand additional renewable energy:** AI companies and data center operators should bear an obligation to develop additional renewable energy capacity in order to offset the extra demand generated by their operations (“additionality”) (Kyriakarakos, 2025; Ebert, Alder et al., 2026). Such a requirement would mitigate the risk that these actors appropriate a disproportionate share of a finite supply of renewable energy. This, in turn, preserves access to renewable energy for other crucial sectors of the economy (e.g., industry, agriculture, transport, housing). The regulatory framework would have to specify clear definitions and criteria that determine which forms of energy consumption fall within the scope of this obligation.

3. **Mandate waste heat recovery for new data centers:** Data centers produce large amounts of heat that can be used productively in industrial facilities or to heat homes, and as such replace fossil fuels. Furthermore, this approach mutually benefits data center operators, as application of waste heat for municipal heating can significantly increase public acceptance of data centers in residential areas (Biermeier & Lange, 2026). Establishing or retaining concrete requirements for waste heat recovery is therefore a major factor for the sustainability of the data center, as included in the German Energy Efficiency Act (EnEfG). Allowing exceptions to such mandates, e.g. based on circumstances of preferred locations, may render such provisions ineffective. In other words, new data centers should ensure that a certain percentage of their waste heat can be used and locations may be constrained accordingly, there should be no exceptions that enable locations where there is no available heating network or off-taker.

4. **Fund development of energy-efficient, specialized compute infrastructure suited to Germany’s industrial needs:** As discussed in Section I, Germany's most important AI use cases do not necessarily require the largest and most power-intensive hardware designed for generative AI workloads. Public investment in energy-efficient, task-specialized compute infrastructure can help Germany secure AI competitiveness while keeping the associated resource footprint within sustainable bounds. This includes support for low-power and energy-efficient compute engines, efficient resource provisioning that accounts for both operational and embodied costs, approaches that extend hardware lifetimes and reduce replacement cycles, and the adoption of emerging sustainable semiconductor and infrastructure technologies. Beyond the design of individual facilities, investment should also support the development of grid-aware data center operations, including techniques that enable data

centers to respond dynamically to real-time grid conditions through intelligent workload scheduling, cooling modulation, and the activation of on-site generation or storage resources. Such approaches can reduce peak electricity demand, smooth load fluctuations, and enable load shifting without compromising the reliability of compute access, transforming data centers from passive electricity consumers into active participants in grid stability. Public-private partnerships are particularly well-suited to driving progress in these areas, enabling German industry and research institutions to develop and deploy sustainable computing solutions that reflect the specific needs of the domestic AI economy.

d) Carbon and electricity pricing

Challenge: Among the potentially most effective ways to address environmental concerns around data centers and AI is to internalize environmental costs, for example in the form of a price on carbon. Generally, carbon pricing is recognized as the leading tool in environmental economics to mitigate climate change and drive a transition to a more sustainable economy (see, e.g., Baranzini et al., 2027; Döbbeling-Hildebrandt et al., 2024; Parry et al., 2022). Especially in a rapidly evolving field such as AI, this approach can be immediately implemented and is a much more feasible option for incentivizing energy efficiency than efficiency standards in model development. Practically, this means that data centers should be priced under existing carbon pricing schemes and exceptions should not be made. There is a concern that data centers are at risk of carbon leakage, for example by moving to areas with no carbon pricing, but a recent first assessment suggests that this risk is small (Schmid et al., 2025). Data centers can pass on costs, have other strong criteria for localization (e.g., data protection requirements), and have historically had a strong interest in sustainability. EU data centers are already powered with 86% renewable energy (European Commission, 2025). Moreover, reducing electricity prices to avoid carbon pricing during a phase of data center buildout might risk inefficient infrastructure for years to come. Economic instruments such as carbon pricing also render many reporting and accounting obligations obsolete, reducing paper work with a more market-driven regime of incorporating sustainability considerations into AI model development and deployment.

Policy recommendations:

1. **Ensure carbon pricing of data centers without compensation:** Ensure that data centers are subject to carbon pricing mechanisms, as that is an effective way to improve the energy efficiency of models and architecture, and promote the use of renewable energy. Exemptions in the interest of localizing datacenters may create the risk of lock-in to inefficient infrastructure, and would result in a missed opportunity to leverage data center owners as investors in renewable energy projects. To avoid carbon prices, grid-connected data centers can easily switch to renewable sources, e.g. by means of power purchase agreements or onsite renewable energy. For Germany and the EU we therefore recommend **no inclusion of data centers in the EU electricity price compensation scheme,** unless they are paired with very stringent efficiency, renewable energy, and waste heat standards. Especially given limited risks for carbon leakage (Schmid et al., 2025), and the grid infrastructure investment needed to facilitate additional data center capacity being passed on to all customers (Jungblut, 2025), compensating electricity prices for data centers could raise fairness concerns. For data

centers building out on-site power generation, there should be no exemptions to carbon pricing, e.g. under the EU Emissions Trading Scheme (EU ETS).

2. **Protecting electricity customers:** While there is no clear evidence yet of electricity price increases due to data center expansion in Germany (Jungblut, 2025), it has been shown for other regions that residential customers face higher prices where there is substantial data center buildout in the region. For example, notable differences in electricity price increases between areas that have many data centers vs. those that do not have been documented in the U.S. (Saul et al., 2025). Residential customers in Ireland, which has among the highest density of data centers in the world, have experienced electricity price increases of 25.9% from 2024 to 2025 (Eurostat, 2025). We recommend that initiatives to rapidly expand data center capacity in Germany include cost-allocation mechanisms to ensure that additional grid infrastructure costs are shouldered by data center owners, rather than ratepayers subsidizing electricity costs for these facilities.

IV. Effects on the Energy Transition

The tech sector and the energy sector are increasingly interdependent, with multifaceted challenges that arise from this development (Weko, 2025). AI and digital service providers act as key suppliers of digital and AI solutions which are increasingly needed in the energy sector (Donti and Kolter, 2021; Hao, 2024) and can be viewed as intangible assets (Weko, 2025). In addition, the sector is also becoming a major electricity customer with the buildout of AI infrastructure (Kamiya & Coroamă, 2025), and technology firms increasingly own generation assets (Cleanview, 2026). There is the concern that AI firms may monopolize their role in supplying digital infrastructure and become a decisive player, and potential risk, for the energy transition. The AI sector's own interest in readily available energy to power new data centers may additionally result in demands for regulatory changes that could potentially slow down the energy transition. Here, we raise two challenges for policy-making: 1) the interest in data center expansion resulting in unsustainable changes to energy regulation, 2) balancing economic interests in the AI infrastructure expansion with sustainability goals.

a) Maintaining fair grid connection and energy market frameworks

Challenge: An interest in fast data center expansion may cause regulatory changes that could slow down the energy transition. For example, modern AI-ready data centers have a large power draw from the grid, and require similar connections as energy storage, renewable energy generation, large customers like industrial plants, charging infrastructure and others. Connecting such a data center to the grid requires significant time and resources. For that reason, there is a long waiting time for grid connections, which is considered a major bottleneck for data center construction and has resulted in many data centers being powered behind the meter (off-grid) in the US (Cleanview, 2026). In Germany, grid operators used to proceed on a first-come-first-served basis, a strategy which is currently being reformed. Changes in the order of serving the connection requests may have profound implications on the energy transition, as many customers competing with the data centers are central for decarbonizing the grid (renewables and electricity storage) and sectors such as transportation and industry (vehicle charging infrastructure and large industrial customers).

Policy recommendation:

1. **Preserve due process and fair procedures, instead of reacting to short-term demand in** regulatory **reform.** It is important not to overrule core energy regulations that are in place to ensure fair and equitable connection decisions as well as a sustainable, safe and economically viable energy system, in favor of fast data center expansion. Regulatory changes with a narrow focus and short time-horizon may lead to long-term problems and a failure to meet energy transition targets. Instead, regulators should tackle issues of grid access more broadly, which would result in additional capacity to serve data center customers and prevent crowding out other much-needed energy infrastructure.

b) Sustainable economic participation in the infrastructure supply chain

Challenge: As German firms supply important products for the global data center expansion, such as gas-fired power plants and generators, Germany may benefit economically from AI infrastructure expansion which may be in tension with sustainability objectives. Particular attention should be paid to the fact that German companies sell energy generation technology to data centers. The supply of natural gas generators for data centers, for example, drive company revenue and stock prices (Millard, 2025). Small gas-fueled generation units on premise are one of the least environmentally sustainable ways to supply data centers with electric power, and those are increasingly being purchased for use behind the meter. Larger, on-grid units, which are also sold by German firms, are more efficient, but they run the risk to prevent the decarbonization of the grid long-term. In the case of German firms providing such units and components, similar economic opportunities may exist with sustainable on- or off-grid renewable and storage installations, but there seems to be less demand in the US market for those solutions (Cleanview, 2026).

Policy recommendations:

1. **Commission a study on German economic interdependencies with AI infrastructure expansion:** We recommend that the German government commissions a study on how German firms benefit from data center construction and operation in Germany and worldwide, what environmental and economic effects relate to that, and how sustainable solutions can be supported. This study can then inform the discussion of how Germany can pioneer specialized sectoral AI products that are beneficial for the energy transition and other aspects of environmental sustainability.
2. **Foster data center localization with high environmental standards:** The demand for AI infrastructure is currently driven by countries such as the US and China, which leave little room for policy intervention from Germany or Europe. This, however, means that fostering data center buildout in Europe with high environmental standards can create a market for innovative products from European companies. If incentives are set right, data center investments can support renewable energy and storage technology industries and foster the local economy, such as with wind energy development.

Conclusion

The rapid integration of artificial intelligence into the economy presents undeniable environmental challenges, from the increasing energy and resource demands of large-scale computation to the complex indirect impacts of its broader deployment. However, as this report underscores, these challenges come alongside significant strategic opportunities. By prioritizing efficient, task-specific systems over resource-intensive general-purpose models, Germany and Europe can transform sustainable AI from a perceived constraint into a core competitive advantage, proving that technological leadership can effectively exist within planetary boundaries.

Realizing this vision requires a robust, forward-looking regulatory structure at the national and international level. Policymakers must ensure that domestic frameworks such as the EU AI Act and the German Energy Efficiency Act (EnEfG) sufficiently capture both the direct computing-related footprint and the broader application-related impacts of AI. Furthermore, AI supply chains and infrastructure are inherently borderless, as are GHG emissions, necessitating cross-border cooperation to develop a global strategy to address the global challenge of sustainable AI.

Finally, the rapid evolution of AI dictates that sustainability cannot be achieved through static, one-off policy interventions; it requires enduring institutional continuity. For instance, **a dedicated Task Force for Sustainable AI in a Competitive Economy** could track technical updates and policy changes in this respect, communicate advancements and challenges to the general public, and inform policy stakeholders about key developments and possible legal adaptations. Only through continuous, adaptable institutional oversight can we ensure that the next generation of AI innovation remains aligned with our long-term environmental stewardship and climate goals.

Annex: Policy Recommendations

I. Beyond Large Language Models and General Purpose AI

1. **Dedicated funding for specialized AI models and use cases** to strengthen the German and European SME-driven AI ecosystem and to promote inherently energy-efficient solutions.
2. **Targeted investment in alternative AI paradigms** beyond current generative models, such as neuro-symbolic AI, Joint-Embedding Predictive Architectures, neuromorphic computing, causal AI, embodied AI and robotics, and probabilistic programming.
3. **Support for an ecosystem of climate-focused AI service providers** with strong domain expertise and transparent cost-benefit profiles, for use cases directly tackling sustainability challenges.
4. **Public-private partnerships** to pool and use industry-scale data within a secure framework (e.g., federated learning), which also unlocks data for specialized models developed for environmentally beneficial tasks.

II. Transparency for Efficient Markets and User Choice

5. **Standardized reporting and benchmarks** that define exactly how environmental metrics should be reported, expand to cover the full model lifecycle, and explicitly include inference-phase impacts.
6. **Broader access to disclosure of impacts** so that companies integrating foundation models into their own software have the data required to calculate and report their own Scope 3 emissions, and researchers and users remain informed.
7. **Independent environmental audits for GPAI models**, which can build on the mandatory third-party audit frameworks established for model safety in the AI Act's GPAI Code of Practice and for data centers in the German Energy Efficiency Act (EnEfG).
8. **Green public procurement policies** at EU and national levels that establish environmental criteria and transparency requirements for AI systems and cloud services.
9. **Transparency on application-related impacts in the AI Act** including, at a minimum, a qualitative description of application-related environmental effects (GHG emissions and water), those arising from how the models are used, for high-risk AI systems.
10. **Fostering impact assessment with public-private partnerships** to pilot AI environmental impact assessments, leveraging existing corporate responsibility frameworks like the Corporate Digital Responsibility (CDR) Initiative to develop scalable, practical reporting methodologies for the broader industry.
11. A **centralized database and regular reporting framework** to track AI use cases over time and to link specific deployments to their resource intensity and societal impacts.

12. **The right to use digital technology without generative AI** via prominent, session-persistent opt-out from AI-generated features, applicable to search engines and other digital services.

13. **The right to use a more sustainable model** by allowing users to default to the most energy-efficient model that still provides the required modality, with reasoning capabilities turned off by default.

III. Sustainable AI Infrastructure

14. **Public disclosure of data center environmental data** with mandated third-party environmental audits that account for large-scale AI workloads, verifiable Scope 2 and Scope 3 GHG emissions, and rated design IT power capacity.

15. **Require rigorous reporting standards for design versus actual operational performance,** for example for the installed information technology power demand reported under the European Energy Efficiency Directive. Operators should be required to report the actual installed electrical design capacity of their IT infrastructure, in addition to actual operational power consumption, to assess utilization rates.

16. **Require granular and dynamic power consumption reporting** to assess how data centers interact with local grids in real time.

17. **Granular and dynamic power consumption reporting** at fine time intervals, with disaggregated figures for training and inference workloads.

18. **Strengthen water reporting to capture consumptive use, peak demand, and local water stress** to reflect real-world operational conditions rather than design estimates

19. **Improve reporting of energy and water consumption under the European Energy Efficiency Directive.** For water, reporting requirements should explicitly define consumption as water lost through evaporation or not returned to its original source. For energy, Cooling Degree Days should be derived automatically from publicly available geospatial weather datasets based on registered facility location.

20. **Support development of AI hardware specific reporting standards by engaging with international efforts** such **as Lot9, NIST, and IEEE** to ensure that AI-specific efficiency metrics are incorporated.

21. **Extend reporting to cover lifecycle carbon and e-waste, with independent validation**; requirements should expand beyond operational energy to include embodied carbon in hardware supply chains, hardware replacement cycles, and e-waste volumes.

22. Going beyond more rigorous reporting standards for design versus actual operational performance (see also Section III b), **efficiency requirements for data centers should be assessed against actual operational performance, not theoretical design scenarios**.

23. **Demand additional renewable energy be incorporated onto grids with new data centers to** mitigate the risk that these actors appropriate a disproportionate share of a finite supply of renewable energy.

24. **Mandate waste heat recovery for new data centers** by establishing or retaining concrete requirements for waste heat recovery is therefore a major factor for the sustainability of the data center, as included in the German Energy Efficiency Act (EnEfG).

25. Public **investment in energy-efficient, task-specialized compute infrastructure** including support for low-power and energy-efficient compute engines, efficient resource provisioning that accounts for both operational and embodied costs, approaches that extend hardware lifetimes and reduce replacement cycles, and the adoption of emerging sustainable semiconductor and infrastructure technologies.

26. **Ensure that data centers are subject to carbon pricing mechanisms** to improve the energy efficiency of models and architecture, and promote the use of renewable energy.

27. **Protecting electricity customers** by requiring initiatives to rapidly expand data center capacity in Germany include cost-allocation mechanisms to ensure that additional grid infrastructure costs are shouldered by data center owners, rather than ratepayers subsidizing electricity costs for these facilities.

IV. Effects on the Energy Transition

28. **Preservation of fair grid connection procedures** – regulators should not overrule core energy regulations in favor of fast data center expansion, as that would risk crowding out renewable energy, storage, and charging infrastructure.

29. **A government-commissioned study** on how German firms benefit from data center construction and operation worldwide, what environmental effects this entails, and how sustainable alternatives can be supported.

30. **Localization with high environmental standards** – if strong standards and mechanisms such as CO_2 pricing are applied, data center construction in Europe becomes not only economically but also environmentally preferable and can support renewable energy and storage technology industries.

References

AAAI (Association for the Advancement of Artificial Intelligence). (2025, March). AAAI 2025 presidential panel on the future of AI research. https://aaai.org/wp-content/uploads/2025/03/AAAI-2025-PresPanel-Report-Digital-3.7.25.pdf

Alder, N., Ebert, K., Herbrich, R., & Hacker, P. (2025). AI, Climate, and Transparency: Operationalising and Improving the AI Act. Journal of European Consumer and Market Law, 14(4), 166-170.

Andrianandrianina Johanesa, T. V., Equeter, L., & Mahmoudi, S. A. (2024). Survey on AI applications for product quality control and predictive maintenance in industry 4.0. Electronics, 13(5), 976.

Baranzini, A., Van den Bergh, J. C., Carattini, S., Howarth, R. B., Padilla, E., & Roca, J. (2017). Carbon pricing in climate policy: seven reasons, complementary instruments, and political economy considerations. Wiley Interdisciplinary Reviews: Climate Change, 8(4), e462.

Belcak, P., Heinrich, G., Diao, S., Fu, Y., Dong, X., Muralidharan, S., ... & Molchanov, P. (2025). Small language models are the future of agentic ai. arXiv preprint arXiv:2506.02153.

Bhushan, C. M., Koppuravuri, P., Prasanthi, N., Gazi, F., Hussain, M. M., Abdussami, M., ... & Faizi, J. (2025). Deploying TinyML for energy-efficient object detection and communication in low-power edge AI systems. Scientific Reports.

Bhuyan, B. P., Ramdane-Cherif, A., Tomar, R., & Singh, T. P. (2024). Neuro-symbolic artificial intelligence: a survey. Neural Computing and Applications, 36(21), 12809-12844.

Biermeier, S., & Lange, J. (2026). Energy playbook for data centers: Building future-proof energy systems for the digital age [White paper]. E.ON Energy Infrastructure Solutions. https://www.eon.com/de/ueber-uns/presse/pressemitteilungen/2026/abwaermenutzung-erhoeht-akzeptanz-von-rechenzentren.html

Bitkom. (2025a). Künstliche Intelligenz in Deutschland: Perspektiven aus Bevölkerung & Unternehmen (Studienbericht). Bitkom e. V.

Bitkom. (2025b, March 27). Industrie 4.0: Schon 42 Prozent der Unternehmen setzen KI in der Produktion ein [Presseinformation]. Bitkom e. V.

Boyd-Madsen, L. (2025). What Ireland's Data Center Crisis Means for the EU's AI Sovereignty Plans. AlgorithmWatch. Retrieved March 25, 2026, from https://algorithmwatch.org/en/ireland-data-center-crisis-eu-sovereignty/

Bundesagentur für Sprunginnovationen. (n.d.). SPRIND Challenge: New Computing. Retrieved 19.03.2026, from https://www.sprind.org/taten/challenges/newcomputing

Bundesministerium für Wirtschaft und Klimaschutz (BMWK). (2025). Stand und Entwicklung des Rechenzentrumsstandorts Deutschland.

https://www.bundeswirtschaftsministerium.de/Redaktion/DE/Publikationen/Technologie/stand-und-entwicklung-des-rechenzentrumsstandorts-deutschland.pdf?__blob=publicationFile&v=1

Cao, J., Ma, Y., Li, X., Ren, Q., & Chen, X. (2026). Task-Specific Efficiency Analysis: When Small Language Models Outperform Large Language Models. arXiv preprint arXiv:2603.21389.

Cleanview. (2026). Bypassing the Grid: How Data Centers Are Building Their Own Power Plants. Retrieved February 4, 2026, from https://cleanview.co/

Clutton-Brock, P., Rolnick, D., Donti, P. L., & Kaack, L. (2021). Climate change and AI. recommendations for government action. GPAI, Climate Change AI, Centre for AI & Climate.

de Melo, G., Koyejo, S., & Mandt, S. (2025). Generative AI and Foundation Models: A Gentle Technical Introduction.

Dehghani, M., Arnab, A., Beyer, L., Vaswani, A., & Tay, Y. (2022). The efficiency misnomer. In 10th International Conference on Learning Representations. https://openreview.net/forum?id=iulEMLYh1uR

Ding, D., Mallick, A., Wang, C., Sim, R., Mukherjee, S., Ruhle, V., ... & Awadallah, A. H. (2024). Hybrid LLM: Cost-efficient and quality-aware query routing. ICLR 2024.

Döbbeling-Hildebrandt, N., Miersch, K., Khanna, T. M., Bachelet, M., Bruns, S. B., Callaghan, M., ... & Minx, J. C. (2024). Systematic review and meta-analysis of ex-post evaluations on the effectiveness of carbon pricing. Nature Communications, 15(1), 4147.

Donti, P. L., & Kolter, J. Z. (2021). Machine Learning for Sustainable Energy Systems. Annual Review of Environment and Resources, 46 (Volume 46, 2021), 719–747. https://doi.org/10.1146/annurev-environ-020220-061831

Donti, P., Rolnick, D., & Kaack, L. (2022) Tutorial: Climate change and machine learning: Opportunities, challenges, and considerations. ICML Tutorial

Ebert, K., Alder, N., Herbrich, R., & Hacker, P. (2026). AI, climate, and regulation: From data centers to the AI Act. Computer Law & Security Review, fothcoming (2026), arXiv preprint arXiv:2410.06681.

Ebert, K., Gamazaychikov, B., Hacker, P., & Luccioni, S. (2026). The Global Landscape of Environmental AI Regulation: From the Cost of Reasoning to a Right to Green AI. arXiv preprint arXiv:2603.00068.

European Commission: Directorate-General for Energy, EY, AIT and Borderstep (2025) Assessment of the energy performance and sustainability of data centres in EU – First technical report, Publications Office of the European Union, https://data.europa.eu/doi/10.2833/3168794

Eurostat (2025). Electricity price statistics https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics

Falk, S., Ekchajzer, D., Pirson, T., Lees-Perasso, E., Wattiez, A., Biber-Freudenberger, L., ... & van Wynsberghe, A. (2025). More than Carbon: Cradle-to-Grave environmental impacts of GenAI training on the Nvidia A100 GPU. arXiv preprint arXiv:2509.00093.

Fernandez, J., Na, C., Tiwari, V., Bisk, Y., Luccioni, S., & Strubell, E. (2025). Energy considerations of large language model inference and efficiency optimizations. Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics. https://arxiv.org/abs/2504.17674

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data?. Advances in neural information processing systems, 35, 507-520.

Gupta, U., Kim, Y. G., Lee, S., Tse, J., Lee, H. H. S., Wei, G. Y., ... & Wu, C. J. (2021, February). Chasing carbon: The elusive environmental footprint of computing. In 2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA) (pp. 854-867). IEEE.

Hacker, P. (2024). Sustainable AI regulation. Common Market Law Review, 61(2).

Hacker, P., Alder, N., Ebert, K., Herbrich, R., Knowles, B., Kaack, L.H., Saucedo, A. (2025) Policy Brief on The EU AI Act for True Environmental Accountability, ACM Europe Technology Policy Committee

Hao, K. (2024, September 13). Microsoft's Hypocrisy on AI. The Atlantic. https://www.theatlantic.com/technology/archive/2024/09/microsoft-ai-oil-contracts/679804/

Heimberger, H., Jäger, A., & Maloča, S. (2024). Künstliche Intelligenz in der Produktion: Digitale Produktionssysteme und Weiterbildung als Voraussetzung für KI? Fraunhofer-Institut für System- und Innovationsforschung ISI.

Hintz, M. J., Gross, M., Creutzig, F., & Kaack, L. H. (2026). Practical implementation of artificial intelligence for climate change mitigation in cities–priorities, collaborations and challenges. Energy Research & Social Science, 131, 104498.

Hintz, M. J., Milojevic-Dupont, N., Creutzig, F., Repke, T., & Kaack, L. H. (2025). A systematic map of machine learning for urban climate change mitigation. Nature Cities, 2(10), 924-936.

Horvat, D., Lerch, C. M., Jäger, A., & Heimberger, H. (2025). Leveraging AI in Production for Product Innovation: Insights from German Manufacturing. In IFIP International Conference on Advances in Production Management Systems (pp. 118-130). Cham: Springer Nature Switzerland.

ICIS. (2025). Data centres: Hungry for power. https://www.icis.com/explore/resources/data-centres-hungry-for-power/

Jeanquartier, F., Jean-Quartier, C., Rieder, P., Misirlić, V., Pasero, C., Hohensinner, R., ... & Holzinger, A. (2026). Assessing the carbon footprint of language models: Towards sustainability in AI. Resources, Conservation and Recycling, 226, 108670.

Jin, Y., Wei, G.-Y., & Brooks, D. (2025). The energy cost of reasoning: Analyzing energy usage in LLMs with test-time compute. arXiv. https://arxiv.org/abs/2505.14733

Jungblut, S.-I. (2025). Germany's Data Center Boom is Pushing the Power Grid to Its Limits. Tech Policy Press. https://techpolicy.press/germanys-data-center-boom-is-pushing-the-power-grid-to-its-limits

Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. Nature Climate Change, 12(6), 518-527.

Kaack, L., Donti, P., Dunietz, J., Klemmer, K., & Milojevic-Dupont, N. (2021, August). Feedback from: Climate Change AI. European Commission - Have Your Say. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2665623_en

Kamiya, G., & Coroamă, V. C. (2025). Data centre energy use: Critical review of models and results. IEA Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E). https://www.iea-4e.org/wp-content/uploads/2025/05/Data-Centre-Energy-Use-Critical-Review-of-Models-and-Results.pdf

Ketan Joshi (2026) Big tech's selective disclosure masks AI's real climate impact. URL https://ketanjoshi.co/2025/08/23/big-techs-selective-disclosure-masks-ais-real-climate-impact/.

Klimareferat der Stadt Frankfurt am Main. (2022). Treibhausgasbilanz 2022 Frankfurt am Main. https://frankfurt.de/-/media/frankfurtde/frankfurt-themen/klima-und-energie/pdf/klimareferat/treibhausgasbilanz/treibhausgasbilanz_2022_frankfurt-am-main.pdf

Krupp, L., Geißler, D., Lukowicz, P., & Karolus, J. (2025). Towards sustainable web agents: A plea for transparency and dedicated metrics for energy consumption. arXiv preprint arXiv:2502.17903.

Kyriakarakos, G. (2025). Bridging the AI–Energy Paradox: A Compute-Additionality Covenant for System Adequacy in Energy Transition. Sustainability, 17(21), 9444.

Lei, N., Lu, J., Shehabi, A., & Masanet, E. (2025). The water use of data center workloads: A review and assessment of key determinants. Resources, Conservation and Recycling, 219, 108310.

Li, P., Yang, J., Islam, M. A., & Ren, S. (2025). Making AI less "thirsty." Communications of the ACM, 68(7), 54–61. https://dl.acm.org/doi/10.1145/3724499

Li, P., Yang, J., Islam, M. A., & Ren, S. (2025). Making ai less' thirsty'. Communications of the ACM, 68(7), 54-61.

Luccioni, A. S., Strubell, E., & Crawford, K. (2025). From efficiency gains to rebound effects: The problem of Jevons' paradox in AI's polarized environmental debate. In Proceedings of the 2025 ACM conference on fairness, accountability, and transparency (pp. 76-88).

Luccioni, S., & Gamazaychikov, B. (2025, December 4). AI Energy Score v2: Refreshed leaderboard, now with reasoning. Hugging Face Blog. https://huggingface.co/blog/sasha/ai-energy-score-v2

Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power hungry processing: Watts driving the cost of AI deployment?. In Proceedings of the 2024 ACM conference on fairness, accountability, and transparency (pp. 85-99).

Massenkoff, M., Lyubich, E., McCrory, P., Appel, R., & Heller, R. (2026). The Anthropic Economic Index report: Learning curves. Anthropic. https://www.anthropic.com/research/economic-index-march-2026-report

Millard, R. (2025). Siemens Energy orders hit record on ‘enormous’ data centre demand. Financial Times. https://www.ft.com/content/f6c20f7a-c9f7-439d-aab7-6c97d049d229

Ngo-Ho, A. K. N., Chauvin, M., Gosset, S., Cordier, P., & Gamazaychikov, B. (2026). Balancing Sustainability and Performance: The Role Of Small-Scale LLMs In Agentic Artificial Intelligence Systems. arXiv preprint arXiv:2601.19311.

NVIDIA. (2025). Product carbon footprint (PCF) summary for NVIDIA HGX H100. https://images.nvidia.com/aem-dam/Solutions/documents/HGX-H100-PCF-Summary.pdf

Onitiu, D., Wachter, S., & Mittelstadt, B. (2026). The fallacy of sustainable Generative AI: limitations in EU environmental regulation of data centres and paths forward, Working Paper, 2026.

Parry, I., Zhunussova, K., & Black, M. S. (2022, July). Carbon taxes or emissions trading systems?: instrument choice and design. IMF Staff Climate Note 2022/006.

Pecher, B., Srba, I., & Bielikova, M. (2025, November). Comparing specialised small and general large language models on text classification: 100 labelled samples to achieve break-even performance. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (pp. 165-184).

Rolnick, D., Aspuru-Guzik, A., Beery, S., Dilkina, B., Donti, P. L., Ghassemi, M., ... & White, A. (2024). Application-driven innovation in machine learning. arXiv preprint arXiv:2403.17381.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... & Bengio, Y. (2022). Tackling climate change with machine learning. ACM Computing Surveys (CSUR), 55(2), 1-96.

Sarker, M. K., Zhou, L., Eberhart, A., & Hitzler, P. (2022). Neuro-symbolic artificial intelligence: Current trends. AI Communications, 34(3), 197-209.

Saul, J., Nicoletti, L., Pogkas, D., Bass, D., & Malik, N. (2025). AI Data Centers Are Sending Power Bills Soaring. Bloomberg.Com. https://www.bloomberg.com/graphics/2025-ai-data-centers-electricity-prices/

Schmid, N., Coroamă, V. C., Dumbravă, o., Eichler, M., Reisser, M., Kaack, L. H., Axenbeck, J., Füssler, J. (2025) Carbon leakage in AI-driven data center growth?. German Environment Agency

Schneider, I., Xu, H., Benecke, S., Patterson, D., Huang, K., Ranganathan, P., & Elsworth, C. (2025). Life-cycle emissions of ai hardware: A cradle-to-grave approach and generational trends. arXiv preprint arXiv:2502.01671.

Schulze, M., Veen V. C (2025). Four Key Conclusions about Data Center Environmental Transparency in the Netherlands. Leitmotiv

Shnitzer, T., Ou, A., Silva, M., Soule, K., Sun, Y., Solomon, J., ... & Yurochkin, M. (2023). Large language model routing with benchmark datasets. arXiv preprint arXiv:2309.15789.

Van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. AI and Ethics, 1(3), 213-218.

Wang, Y., Chen, W., Han, X., Lin, X., Zhao, H., Liu, Y., ... & Yang, H. (2024). Exploring the reasoning abilities of multimodal large language models (mllms): A comprehensive survey on emerging trends in multimodal reasoning. arXiv preprint arXiv:2401.06805.

Weko, S. (2025). New Sites of Accumulation? Why Intangible Assets Matter for Energy Transitions. Review of Political Economy, 37(4), 1571–1598. https://doi.org/10.1080/09538259.2024.2391304

Wu, C. J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., ... & Hazelwood, K. (2022). Sustainable ai: Environmental implications, challenges and opportunities. Proceedings of machine learning and systems, 4, 795-813.

Yang, C., Zhao, R., Liu, Y., & Jiang, L. (2025). Survey of specialized large language model. arXiv preprint arXiv:2508.19667.

Yu, F., Zhang, H., Tiwari, P., & Wang, B. (2024). Natural language reasoning, a survey. ACM Computing Surveys, 56(12), 1-39.