AI, Climate, and Transparency: Operationalizing and Improving the AI Act

Nicolas Alder[†]
Hasso Plattner Institute
Potsdam, Germany
nicolas.alder@hpi.de

Kai Ebert[†]
European University Viadrina
Frankfurt (Oder), Germany
kebert@europa-uni.de

Ralf Herbrich
Hasso Plattner Institute
Potsdam, Germany
ralf.herbrich@hpi.de

Philipp Hacker
European University Viadrina
Frankfurt (Oder), Germany
hacker@europa-uni.de

Abstract

This paper critically examines the AI Act's provisions on climate-related transparency, highlighting significant gaps and challenges in its implementation. We identify key shortcomings, including the exclusion of energy consumption during AI inference, the lack of coverage for indirect greenhouse gas emissions from AI applications, and the lack of standard reporting methodology. The paper proposes a novel interpretation to bring inference-related energy use back within the Act's scope and advocates for public access to climate-related disclosures to foster market accountability and public scrutiny. Cumulative server level energy reporting is recommended as the most suitable method. We also suggest broader policy changes, including sustainability risk assessments and renewable energy targets, to better address AI's environmental impact through transparency measures that facilitate competition among model providers and allow conscious consumer decisions. Finally, we provide first insights on the General-Purpose AI Model Code of Practice as it relates to climate reporting.

1 Introduction

The climate implications of artificial intelligence (AI), including energy and water consumption, are increasingly subjected to public scrutiny and academic research [Strubell et al., 2020, Dodge et al., 2022, Kaack et al., 2022, Luccioni et al., 2023, Li et al., 2025, Hacker, 2024]. In one respect, the application of AI in environmental issues provides numerous promising opportunities for more effectively tackling climate and sustainability challenges [Rolnick et al., 2022, Wu et al., 2022, Cowls et al., 2023]. These include advancements in carbon measurement for cloud computing [Dodge et al., 2022], utilizing earth observation to aid electricity grids in achieving carbon neutrality [Persello et al., 2022, Chandra et al., 2025], enhanced decision-making for carbon capture technologies [Chandra et al., 2025], and accelerating scientific processes across various fields [Wang et al., 2023].

In another respect, energy efficiency targets for data centers are under discussion,[1] and there is concern that their energy consumption could surpass the available supply of renewable energy. Major

[†]Equal contribution.

[1]Further EU legislation can be expected following the Commission's report under Art. 12(5) Energy Efficiency Directive (EU) 2023/1791 and Commissioner Dan Jørgensen announcements of a data center energy efficiency package at the IEA 10th Annual GlobalConference on Energy Efficiency on 12 June 2025.

Workshop on Regulatable ML at the 39th Conference on Neural Information Processing Systems (NeurIPS 2025).

companies like Google have reported that increased energy demand related to AI endangers their carbon zero strategies [Google, 2024].

As in many collective action problems, regulation may play a major part in mitigating the negative impact of AI on climate while fostering socially beneficial use cases. Additionally, recent studies suggest that thoughtful regulation even fosters necessary green innovation in the industry [Zhang et al., 2024]. Globally, several initiatives are underway to establish legal frameworks for AI. The most prominent example is probably the EU AI Act,[2] which has just become applicable at the beginning of August 2024. The Act also includes significant sections concerning climate impacts, primarily reporting obligations. Thus, one might hope, the climate effects of AI could become a relevant market parameter; have reputational repercussions; and enable public scrutiny, by analysts and NGOs.

Against this background, this article offers a novel legal-technical analysis of the transparency provisions of the AI Act, making three core contributions. First, we demonstrate that the Act falls short in several critical areas. Second, we show that even within its current scope, operationalizing its mandates poses significant challenges. Third, we propose targeted reforms for legislators: (i) reinterpret Art. 53 and Annexes XI and XII to encompass inference energy; (ii) extend reporting to indirect greenhouse gas emissions and data-center water use; (iii) revoke the exemption from climate reporting obligations for open-source models; (iv) operationalize energy disclosure via cumulative server-level measurement and provide PUE numbers; (v) mandate public release of all climate-related data. These recommendations would provide the Act with the rigor and accountability essential for sustainable AI governance and inform broader policy initiatives, including the United Nations' Global Digital Compact currently under discussion at the United Nations.

2 Climate Transparency and the AI Act: Gaps and Interpretation Challenges

Any change for the better starts with information about what is wrong. However, at the moment, it is often unclear what the exact impact of the development and usage of an AI model is concerning energy and water consumption. The AI Act seeks to provide a remedy by forcing certain AI providers to make climate-related disclosures. However, the patchwork of provisions includes seven significant ambiguities and loopholes.

First, for high-risk AI systems, providers are required under Art. 11(1) to document the computational resources used in development, training, testing, and validation, as per Annex IV(2). However, there is no explicit requirement to disclose energy consumption, limiting the comparability of, and transparency on, the environmental impact of these high-risk systems to estimates based on the documented computational resources.

Second, the AI Act imposes transparency obligations on providers of general-purpose AI (GPAI) models, particularly concerning energy consumption. Under Art. 53(1)(a), providers must maintain up-to-date technical documentation that includes information specified in Annex XI, which requires known or estimated energy consumption of the model, with estimates potentially based on computational resources. However, this requirement focuses on the model's development phase, excluding the inference phase, as it forms part of the requirements in Annex XI Section 1 para. 2, which refers to "information of the process for the development [of the model]." If indeed the energy consumption is restricted to training and excludes inference (see below, V.), this constitutes a significant oversight given the potentially much greater cumulative energy consumption during inference [Luccioni et al., 2024]. To address this gap, a novel interpretation can be considered. Art. 53(1)(a) and (b), in conjunction with Annex XI and Annex XII, require providers to include in the documentation for downstream AI system providers and authorities information on the technical means needed to integrate the GPAI model into AI systems. Although energy consumption is not explicitly mentioned, these provisions should, arguably, be interpreted to include information on hardware requirements, allowing downstream providers to estimate the energy consumption for inference. This novel interpretation would indirectly ensure transparency regarding inference energy use.

[2]Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 [2024] OJ L 2024/1689.

A third issue arises with open-source (OS) GPAI models, which are generally exempt from transparency obligations unless they pose a systemic risk (Art. 53(2)). Recital 102 emphasizes transparency for OS models but does not include energy consumption in the information that must be disclosed. Rather, the focus is on parameters, model architecture, and usage information, leaving a gap in transparency regarding the energy impact of these models.

Fourth, the AI Act overlooks the greenhouse gas (GHG) effects of AI applications, such as those used in oil and gas exploration. This omission leaves a significant gap, as these applications can substantially contribute to climate change, yet their environmental impact remains unreported. Fifth, while the Act requires energy consumption to be documented, this information is only available to authorities, not downstream providers (unless our suggested interpretation is adopted), and certainly not to the general public. Without broader access to this data, transparency and accountability are significantly curtailed, hindering market effects based on climate reporting, independent research and verification, and public scrutiny by analysts and NGOs. Finally, the Act also fails to address the use of toxic materials and water consumption, a critical factor in data center operations. While most data centers in the EU must report their water usage under the Energy Efficiency Directive, the AI Act lacks a specific attribution to AI, as stipulated for energy consumption, and computing outside the EU is not covered. Given the significant water usage for cooling in data centers, this omission leaves a major aspect of AI's environmental impact unreported.

3 Operationalizing the Requirements: Implementation Challenges

As the previous section showed, under the current version of the AI Act, GPAI providers must log the energy consumption used for training GPAI models. To operationalize this provision, it is crucial to clarify how energy consumption should be measured or estimated. We discuss three methods: measurement at the data center level; at the cumulative server level; and at the individual graphic-processing unit (GPU) level.

Energy efficiency in data centers is measured by the Power Usage Effectiveness (PUE) metric. It denotes the ratio of total energy used by the data center to the energy consumed by its computational hardware. A lower PUE indicates higher energy efficiency, with a global average PUE of 1.58 recorded in 2023 [Statista, 2023]. When measuring energy consumption at the data center level, the advantage lies in capturing the total power usage, including both direct computing energy and overhead like cooling. This provides a comprehensive overview and encourages efficient data center selection. However, it can obscure the energy impacts of specific model architecture or software inefficiencies, as these are influenced by the data center's overall efficiency. Estimating with the PUE ratio is practical but may lack precision for specific model-level insights.

At the cumulative server level, i.e., for all utilized servers within one data center, energy measurement with power distribution units is highly accurate, closely reflecting model size, data volume, and software efficiency. This method is recognized in the industry and can provide detailed insights into energy consumption. However, not all data centers currently track power demand at this level, and implementing such systems can be time-consuming [The Green Grid, 2023]. While cloud providers like AWS and Azure may have these capabilities, widespread reporting standards are lacking, potentially disadvantaging smaller companies.

Finally, measuring energy usage at the GPU level within a server is straightforward with on-chip sensors for components like NVIDIA GPUs, which offer user-friendly monitoring. However, this approach significantly underestimates total energy consumption as it only accounts for a single component, missing the broader picture of server-wide energy use. Therefore, it is not recommended for comprehensive energy tracking.

4 Discussion and Policy Proposals

The AI Act is a first step toward mandatory AI-related climate reporting, but is riddled with loopholes and vague formulations. To remedy this, we make five key policy proposals. Such mechanisms should not only be included in the evaluation report due in August 2028 (Art. 111(6)), but in any interpretive guidelines by the AI Office and other agencies, reviews and potential textual revisions beforehand.

The primary weakness of the AI Act is the exclusion of inferences from explicit and mandatory energy consumption reporting. While we offer a solution for interpretation, it is unclear whether courts, agencies and companies will follow this route. This significantly hampers the assessment of future AI energy usage, related carbon emissions, and effects on (renewable) energy infrastructure. Hence, future guidance from the AI Office, and delegated acts by the Commission (Art. 53(5) and (6)), should explicitly include inference as a reporting category, both in Annex XI (for the AI office) and XII (for downstream actors).

Another major challenge is the failure to include indirect emissions by AI applications (e.g., for oil and gas exploration) and water consumption within the reporting obligations. This should be remedied at the provider (water) and the deployer level (applications).

Table 1: Shortcomings in the AI Act Concerning Climate Reporting, and Policy Proposals

Shortcomings	Policy Proposals
1. Inference Energy Consumption Exclusion	Explicitly include inference in energy reporting obligations in Annexes XI and XII.
2. Indirect Emissions and Water Consumption	Extend reporting obligations to include water consumption and indirect GHG emissions from AI applications.
3. Open-Source Models	Revoke the exemption to ensure comprehensive climate reporting.
4. Lack of Standard Reporting Methodology	Measure energy consumption at the cumulative server level, with separate PUE.
5. Lack of Public Access to Energy Data	Make all climate-related disclosures publicly available to foster transparency and market accountability.

The open source exemption, third, should be revoked. There is no convincing reason to abstain from climate reporting only because other parts of the model are made public and transparent.

Fourth, energy consumption measurements ought to be conducted at the cumulative server level and reported accordingly. This reflects the total computation-related power usage. Furthermore, the PUE factor of each data center, as measured and reported under the Energy Efficiency Directive (EU) 2023/1791 and Delegated Regulation (EU) 2024/1364,[3] provides information for a relevant estimate of the overall energy consumption. By reporting these numbers separately, we can differentiate between the power usage specific to the model (server-level computation) and the efficiency of the data center, thus reflecting the realistic overall energy investment. Estimations for server-level power consumption should utilize peak utilization values from the hardware manufacturer (e.g., NVIDIA). When actual measurements are available, they must be prioritized over estimations. The ultimate aim is to secure as precise power consumption data as possible, allowing for flexibility for model providers with limited access to data infrastructure (such as for finetuning with substantial modification), while also ensuring that estimations are not misused to avoid accurate measurement reporting. These considerations should inform both the technical standards drafted under Art. 40 AI Act and the possible implementation of the Global Digital Compact at the international level.

Finally, fith, all climate-related disclosures must urgently be made available to the general public, not only to authorities and, potentially, downstream actors. Trade secrets and intellectual property do not stand in the way if only aggregate numbers at the cumulative server level are reported. Only in this way, market pressure can build up, reputational effects set in, and public scrutiny via analysts, academics, and NGOs unfold its incentivizing force.

5 The Code of Practice

Importantly, however, the AI Act rules on climate effects have just received a major update as they constitute one part of the recently published General Purpose AI Code of Practice (CoP) [European Commission, 2025a]. It complements the AI Act by offering a voluntary, yet concrete framework for providers of general purpose AI models to demonstrate compliance with the regulatory obligations in the GPAI chapter of the AI Act. The Code emerged from a transparent, multi stakeholder drafting effort led by the European AI Office and independent researchers with participation from nearly 1,000

[3]Commission Delegated Regulation (EU) 2024/1364 on the first phase of the establishment of a common Union rating scheme for data centres [2024] OJ L 2024/1364.

experts across industry, civil society, academia, rights holders, and Member State representatives. It addresses three core chapters—Transparency, Copyright, and Safety & Security—each mapping to corresponding obligations under Articles 53 and 55 of the AI Act.

The Transparency chapter provides a user friendly Model Documentation Form to aid providers in meeting the AI Act's requirements for technical documentation, copyright information, and training data summaries under Article 53. The Copyright chapter guides providers through implementing policies necessary to comply with EU copyright law obligations under the same article. The Safety & Security chapter applies only to providers of the most advanced, high risk models—those subject to Article 55, and offers concrete measures for systemic risk taxonomy, risk assessment, incident reporting, and cybersecurity.

The final version of the Code was published on 10 July 2025 and approved by the Commission and the AI Board on 1 Aug 2025 [European Commission, 2025b]. It will serve as a practical early implementation tool to ease compliance burden. Signatories will benefit from a presumption of conformity (Art. 53(4) and 55(2) AI Act). A braod number of companies, including OpenAI, have since signed and committed to abide by the CoP [European Commission, 2025a].

The CoP, in its Transparency chapter, introduces reporting requirements on energy consumption—notably during both the training and inference phases of GPAI models [Oliver and Bommasani, 2025]. Providers must disclose the amount of energy used for training, expressed in megawatthours (MWh) and recorded with at least two significant figures (e.g. 1.0×10^2 MWh).

In the absence of a delegated act pursuant to Article 53(5) of the AI Act specifying standardized measurement methodologies, providers must describe the method used to measure or estimate training energy consumption. This includes explaining whether the estimation is based on direct measurements or indirect calculations derived from known computational resources. If critical data are unavailable, providers must identify the specific information they lack.

The real surprise comes in the section on inference. Published after the initial publication of this paper as a Working Paper, the CoP takes up our general suggestion to interpret the AI Act in a way compatible with inference reporting, but based on a different legal reasoning. The CoP requires disclosure of the benchmarked amount of computation, reported in floating point operations and recorded with at least two significant figures (e.g. 5.1×10^{17} FLOPs). Providers must describe the task used to benchmark inference computation, such as generating 100,000 tokens, and detail the hardware setup, for example using 64 Nvidia A100 GPUs. As a basis in the AI Act, the CoP cites Annex XI Section 1 para. (2)(e), which requires the disclosure of the "known or estimated energy consumption of the model" [Oliver and Bommasani, 2025]. Indeed, this requirement, read on its own, does not limit the consumption to training. However, as mentioned (Part II.), the heading of Section 1 para. 2 suggests that the item indeed may refer to training only ("development"). If this reading was adopted, then our interpretation provides a further justification for keeping inference in the CoP, and requiring reporting more generally from GPAI providers: it forms part of the "technical means needed to integrate the GPAI model into AI systems" (Annex XI Section 1 para. (2)(a), Annex XII para. (2)(a)). This reading has the additional benefit that the benchmarked inference reporting must be disclosed not only to the AI Office and National Authorities (Annex XI), but also to downstream deployers (Annex XII). The lack of access for the general public, however, remains problematic.

Finally, we do note, however, that some information on energy consumption may ultimately be accessed by researchers. While the public, summarized versions of the Framework and Model Reports (Measure 10.2, Safety and Security Chapter) will likely not contain any information on energy, independent evaluators receive free access to the most capable model versions to facilitate post-market monitoring (Measure 3.4, Safety and Security Chapter). This not only mirrors, to a certain extent, Art. 40 DSA, but also provides researchers with some access to models—and potentially their inference calculations as it can be argued that environmental effects do form part of systemic risks [Kasirzadeh et al., 2025]. However, the measure only applies, like the entire Safety and Security Chapter, only to GPAI models with systemic risk, which limits its effectiveness in tracking model inference energy consumption.

Overall, the CoP does address the problem of inferences in a laudable way; and it provides some limited external oversight via researchers. But generally, the policy suggestions we have sketched remain to be enacted, through guidelines, implementing acts, a revision of the Code of Practice, or even the AI Act itself.

6 Conclusion

This paper tackles some of the complexities at the intersection of AI, climate and regulation. The AI Act does contain significant climate reporting obligations. By drawing on technical and legal research, we show that they contain too many loopholes, and are difficult to operationalize. Perhaps most importantly, even though recent research has shown inference to be a major driver of AI-related GHG emissions, this key area is omitted from the AI Act. A novel interpretation of the Act's reporting obligations might bring inference back within its scope. Furthermore, none of the climate disclosures are initially open to the public. We suggest changing this urgently to kickstart market pressure, induce reputational effects among consumers, and enable crucial public scrutiny, e.g. by academics and NGOs, beyond Measure 3.4 of the CoP's Safety and Security Chapter. For model providers, our transparency proposals enable essential global competition for green innovation in AI models, which currently exists through very indirect and intransparent economic measures, such as electricity prices or (local) water regulations. For consumers, both organizations and individual end-users, this allows for more informed decisions from both environmental and reputational standpoints. Additionally, downstream consumers can select the most cost-effective model for their specific use cases, as inference energy is directly linked to power consumption and hardware expenses.

However, climate reporting can only be a first step in addressing the massive and fast-rising environmental impact of AI models and systems. It must be complemented by substantive obligations, including sustainability risk assessment and management, renewable energy targets for data centers, and potentially even (tradable) caps on the energy and water consumption of data centers and similar major consumption drivers in the AI value chain [Hacker, 2024].

References

Emma Strubell, Ananya Ganesh, and Andrew McCallum. Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 13693–13696, 2020.

Jesse Dodge, Taylor Prewitt, Remi Tachet des Combes, Erika Odmark, Roy Schwartz, Emma Strubell, Alexandra Sasha Luccioni, Noah A Smith, Nicole DeCario, and Will Buchanan. Measuring the carbon intensity of ai in cloud instances. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*, pages 1877–1894, 2022.

Lynn H Kaack, Priya L Donti, Emma Strubell, George Kamiya, Felix Creutzig, and David Rolnick. Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12(6): 518–527, 2022.

Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat. Estimating the carbon footprint of bloom, a 176b parameter language model. *Journal of machine learning research*, 24(253):1–15, 2023.

Pengfei Li, Jianyi Yang, Mohammad A Islam, and Shaolei Ren. Making ai less' thirsty'. *Communications of the ACM*, 68(7):54–61, 2025.

Philipp Hacker. Sustainable ai regulation. *Common Market Law Review*, 61:345–386, 2024.

David Rolnick, Priya L Donti, Lynn H Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, et al. Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)*, 55(2):1–96, 2022.

Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga, Jinshi Huang, Charles Bai, et al. Sustainable ai: Environmental implications, challenges and opportunities. *Proceedings of machine learning and systems*, 4: 795–813, 2022.

Josh Cowls, Andreas Tsamados, Mariarosaria Taddeo, and Luciano Floridi. The ai gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *Ai & Society*, 38(1):283–307, 2023.

Claudio Persello, Jan Dirk Wegner, Ronny Hänsch, Devis Tuia, Pedram Ghamisi, Mila Koeva, and Gustau Camps-Valls. Deep learning and earth observation to support the sustainable development goals: Current approaches, open challenges, and future opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2):172–200, 2022.

Anirban Chandra, Marius Koch, Suraj Pawar, Aniruddha Panda, Kamyar Azizzadenesheli, Jeroen Snippe, Faruk O Alpak, Farah Hariri, Clement Etienam, Pandu Devarakota, et al. Fourier neural operator based surrogates for *co_2* storage in realistic geologies. *arXiv preprint arXiv:2503.11031*, 2025.

Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023.

Google. Google 2024: Environmental report. `https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf`, 2024. Accessed: 06-Jun-2025.

Wanli Zhang, Bin Zhu, Yongling Li, and Dan Yan. Revisiting the porter hypothesis: a multi-country meta-analysis of the relationship between environmental regulation and green innovation. *Humanities and Social Sciences Communications*, 11(1):1–15, 2024.

Sasha Luccioni, Yacine Jernite, and Emma Strubell. Power hungry processing: Watts driving the cost of ai deployment? In *Proceedings of the 2024 ACM conference on fairness, accountability, and transparency*, pages 85–99, 2024.

Statista. What is the average annual power usage effectiveness (pue) for your largest data center? `https://www.statista.com/statistics/1229367/data-center-average-annual-pue-worldwide/`, 2023. Accessed: 18-Aug-2025.

The Green Grid. It & power efficiency survey results (for euc). `https://www.thegreengrid.org/en/resources/library-and-tools/572-IT-%26-Power-Efficiency-Survey-Results-%28for-EUC%29`, 2023. Accessed: 06-Jun-2025.

European Commission. The general-purpose ai code of practice. `https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai`, 2025a. Accessed: 15-Jul-2025.

European Commission. Drawing-up a general-purpose ai code of practice. `https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice`, 2025b. Accessed: 21-Aug-2025.

Nuria Oliver and Rishi Bommasani. Code of practice for general-purpose ai models. transparency chapter. `https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai`, 2025. Accessed: 19-Aug-2025.

Atoosa Kasirzadeh et al. Too big to regulate? systemic risk in market, ai, and platform regulation. Working Paper, forthcoming, 2025.