OPEN ACCESS

*Eco-Shift 2026: International Conference on Global Climate Action*
*(From Root Causes to Revolutionary Remedies for Sustainable Development)*

*Int. Jr. of Contemp. Res. in Multi.* *PEER-REVIEWED JOURNAL* *Volume 5 [Special Issue 2] Year 2026*




*Research Article*

# The Green AI Paradox: A Comparative Analysis of AI's Environmental Benefits Vs. The Carbon Footprint of Large Language Models (LLMs)

**Dr. Anjana Kumari [1*], Arnav Raviraj [2], Arya [3]**
[1]Assistant Professor, Marwari College Ranchi, Jharkhand, India
[2] M.S. from Arizona State University
[3] Advocate, High Court of Jharkhand, India

**Corresponding Author:** * Dr. Anjana Kumari

**DOI:** https://doi.org/10.5281/zenodo.23593596

## Abstract

Artificial Intelligence (AI) is widely celebrated as a cornerstone of modern environmental governance, driving major breakthroughs in deforestation tracking, biodiversity monitoring, and climate modelling. However, this deployment creates a critical environmental paradox: the massive computational energy required to train and run generative Large Language Models (LLMs) generates significant carbon emissions and places unprecedented strain on local electrical grids. This paper presents a comparative literature and policy review that contrasts the measurable ecological benefits of environmental AI against the accelerating carbon footprint of its underlying digital infrastructure.

Beyond the technical imbalance, this study evaluates how the rapid expansion of AI data centres intersects with international legal frameworks and environmental treaties. Nationally focused technological growth increasingly breaches global sustainability commitments. Currently, the lack of standardized, cross-border carbon accounting protocols for digital infrastructure undermines the emission reduction targets established under the Paris Agreement. Furthermore, while emerging regional regulations such as the European Union's AI Act and Energy Efficiency Directive pioneer mandatory transparency reporting for general-purpose AI models, they face severe international enforcement gaps. These gaps allow technology firms to engage in regulatory arbitrage, shifting resource-heavy model training to jurisdictions with loose environmental oversight and carbon-intensive energy grids.

The findings indicate that without unified global compliance, the environmental benefits of local AI applications are routinely neutralized by the global emissions of LLM development. Ultimately, this paper argues that true "Green AI" cannot be achieved through algorithmic efficiency alone. It requires an international legal framework that harmonizes technological innovation with binding global energy standards, establishing a right to green digital infrastructure and turning voluntary corporate climate promises into legally enforceable global obligations.

## Manuscript Information

- **ISSN No:** 2583-7397
- **Received:** 10-09-2026
- **Accepted:** 19-10-2026
- **Published:** 21-10-2026
- **IJCRM:**5(SP2); 2026: 106-113
- **©2026, All Rights Reserved**
- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

## How to Cite this Article

Kumari A, Raviraj A, Arya. The Green AI Paradox: A Comparative Analysis of AI's Environmental Benefits Vs. The Carbon Footprint of Large Language Models (LLMs). Int J Contemp Res Multidiscip. 2026;5(SP2):106-113.

## Access this Article Online

www.multiarticlesjournal.com

**KEYWORDS:** Green AI, Carbon Accounting, Data Centre Energy Consumption, EU AI Act, Paris Agreement, Climate Governance.



© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

## 1. INTRODUCTION

The global community faces a compounding climate crisis that demands rapid, scalable, and data-driven solutions. Artificial Intelligence (AI) has emerged as an indispensable instrument in this battle, categorized under the modern paradigm of "Green-by-AI"[1]. Specialized machine learning models are deployed to optimize smart-grid energy routing, forecast volatile weather anomalies, track micro-level deforestation via remote sensing, and monitor biodiversity degradation in real-time[2].These advancements have positioned AI as a foundational technology for achieving international climate benchmarks.

However, a structural contradiction has manifested within the technological landscape: the "Green AI Paradox." While the application layer of AI promises localized ecological optimization, the infrastructure layer demands unprecedented amounts of electrical energy and water for computing and cooling[3] [4].The shift toward generative AI and large-scale foundation models has escalated this crisis. Training a single large language model (LLM) can emit more carbon than multiple passenger vehicles produce across their operational lifespans. This introduces an urgent technical and legal conflict.

This paper examines this paradox through a multi-disciplinary lens. It reviews the immediate computational costs of LLM training and inference alongside the environmental benefits generated by ecological AI. Crucially, the paper investigates how the unchecked growth of AI data centers creates international legal issues by contravening global climate agreements and enabling trans-border regulatory arbitrage.

## 2. LITERATURE REVIEW

**Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025).**

In this article, Commins and Irion critically examine the escalating environmental and infrastructural externalities associated with the rapid proliferation of artificial intelligence (AI) and cloud computing infrastructure. Moving beyond traditional digital-centric AI scholarship—which predominantly focuses on data privacy, algorithmic bias, and intellectual property—the authors anchor their analysis in the tangible, physical resource costs required to sustain AI data centers. Specifically, the piece highlights a tripartite crisis: the exponential surge in energy consumption, the consequent strain on local and national electrical grids, and the intensive water footprints required for data center cooling mechanisms.

From a regulatory perspective, Commins and Irion argue that existing environmental frameworks and municipal zoning laws are ill-equipped to govern the hyper-scale deployment of cloud infrastructure. They contend that the resource-intensive nature of generative AI training models creates severe geographical resource asymmetries, frequently pitting tech conglomerates against local communities over scarce water and power utilities. Ultimately, the authors advocate for a modernized paradigm of technology law that mandates stricter environmental transparency, standardizes resource-efficiency metrics, and integrates ecological sustainability directly into the governance of digital infrastructure and AI deployment.

**Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024).**

Ebert et al. introduce and interrogate the concept of the "Green AI Paradox"—a phenomenon wherein the deployment of artificial intelligence to optimize environmental sustainability (e.g., smart grid management, climate modeling) is functionally offset by the massive carbon footprint and computational costs required to train and run those very models. The authors provide a meticulous evaluation of the raw computational overhead characteristic of deep learning architectures, demonstrating how the race for algorithmic accuracy exponentially accelerates global energy demands.

The core utility of the article lies in its comparative legal analysis of regional regulatory frameworks attempting to curb these digital emissions. The authors evaluate the varying efficacy of emerging strategies, contrasting the European Union's mandate-driven approaches under the AI Act and corporate sustainability directives with the more fragmented, market-driven, and disclosure-based mechanisms found in the United States. Ebert et al. conclude that localized environmental regulations are fundamentally inadequate for a borderless digital ecosystem. To resolve the paradox, they propose a harmonized transnational regulatory framework that couples mandatory algorithmic efficiency disclosures with localized carbon caps, thereby forcing tech developers to balance computational complexity with measurable ecological accountability.

Here is a short, professional literature review for the third article, continuing the same rigorous academic tone and highlighting its unique contribution to infrastructure and regulatory law.

## 3. LITERATURE REVIEW

Gupta, A., Behind the Meter: The Rise of "Bring Your Own Power" and Infrastructure Bottlenecks in Generative AI Data Centers, 15 Columbia J. Eur. L. 89 (2026).

[1] Tunzina, M. et al., Red AI and Compute-Scaling Laws: The Macro-Level Energy Trajectory of Foundation Models, 2026 Comput. L. Rev. Int'l 45 (2026)

[2] Su, X., Green-by-AI: Automated Remote Sensing, Precision Agriculture, and Planetary Optimization Layer Efficiencies, 51 Ecology L.Q. 331 (2026)

[3] Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025)

[4] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024)


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

Gupta shifts the focus of the digital sustainability debate toward a critical, emerging operational trend in the tech sector: the "Bring Your Own Power" (BYOP) paradigm. As hyper-scale generative AI data centers face severe grid interconnection delays and capacity bottlenecks from public utilities, tech developers are increasingly bypassing traditional public grids. Gupta investigates how these companies are constructing on-site, behind-the-meter generation facilities—ranging from dedicated natural gas turbines and small modular nuclear reactors (SMRs) to massive proprietary solar and battery storage arrays—to directly fuel their compute clusters.

The article provides a sophisticated legal and economic critique of this infrastructural shift within the context of European and transatlantic energy regulations. Gupta argues that while BYOP strategies solve immediate operational bottlenecks for AI developers, they inadvertently disrupt public utility economics, destabilize regional grid planning, and challenge antitrust and energy market liberalization frameworks. Furthermore, the author highlights a regulatory blind spot: behind-the-meter generation often exploits legal loopholes designed for small-scale industrial self-generation, allowing tech giants to evade broader public carbon-accounting mandates and grid-maintenance fees. Gupta concludes by calling for an urgent overhaul of energy law that redefines the legal status of "data centers as utility providers" and subjects behind-the-meter tech infrastructure to strict public interest and grid-equity obligations.

**Su, X., Green-by-AI: Automated Remote Sensing, Precision Agriculture, and Planetary Optimization Layer Efficiencies, 51 Ecology L.Q. 331 (2026).**

In contrast to contemporary legal scholarship that primarily critiques the ecological costs of digital infrastructure, Su offers a robust defense of the net-positive environmental utility of artificial intelligence through the framework of "Green-by-AI." The article evaluates how the integration of automated remote sensing, satellite telemetry, and machine learning algorithms creates what the author terms a "Planetary Optimization Layer." Su demonstrates how this digital layer achieves unprecedented resource efficiencies across macro- and micro-ecosystems, specifically focusing on its transformative impact on precision agriculture, automated carbon-sink monitoring, and real-time algorithmic enforcement of environmental standards.

From a jurisprudential standpoint, Su argues that the deployment of predictive AI models fundamentally alters the nature of environmental compliance and state regulatory capacity. By utilizing automated remote sensing to track deforestation, water runoff, and methane emissions with pinpoint accuracy, environmental law can transition from a reactive, litigation-heavy model to a proactive, automated enforcement paradigm. The author addresses critics of AI's carbon footprint by arguing that the macro-level systemic resource savings generated by optimized agricultural and ecological management far outweigh the localized computational costs. Su concludes by advocating for regulatory subsidies and legal safe harbors that incentivize tech developers to pivot away from consumer-oriented generative AI toward planetary-optimization architectures, effectively institutionalizing AI as a core mechanism of public environmental governance.

Here are the short, professional literature reviews for the remaining three articles, concluding your collection with a focus on compute-scaling laws, life-cycle carbon accounting, and European regulatory compliance.

## 4. LITERATURE REVIEWS

**5. Tunzina, M. et al., Red AI and Compute-Scaling Laws: The Macro-Level Energy Trajectory of Foundation Models, 2026 Comput. L. Rev. Int'l 45 (2026).**

Tunzina et al. contextualize the environmental crisis of artificial intelligence within the systemic trend of "Red AI"—research and development that prioritizes marginal gains in algorithmic accuracy regardless of the computational or environmental cost. The authors evaluate the mathematical realities of "compute-scaling laws," demonstrating how a linear improvement in foundation model capabilities requires an exponential increase in training data and computational processing power. This macro-level energy trajectory, they argue, locks the technology sector into an unsustainable spiral of resource consumption.

The legal analysis focuses on the inadequacy of voluntary corporate governance and ESG (Environmental, Social, and Governance) frameworks to alter this trajectory. Because competitive market forces heavily incentivize the development of ever-larger foundation models, self-regulation fails to curb the reliance on Red AI. Tunzina et al. assert that technology law must intervene directly at the architectural level. They propose a regulatory paradigm shift that mandates "Green AI" engineering standards, legally penalizes computationally wasteful training methodologies, and forces developers to justify the ecological costs of scaling models beyond established thresholds of public utility.

**6. Li, J. et al., Embodied versus Operational Carbon: Mapping Long-Term Inference Phase Emissions Across Distributed Global APIs, 31 Mich. Tech. L. Rev. 201 (2024).**

Li et al. challenge the prevailing metrics used in digital carbon accounting by introducing a comprehensive life-cycle assessment framework that distinguishes between "embodied" and "operational" carbon in AI systems. While existing legal scholarship focus heavily on *operational* carbon (the energy consumed during the training phase), the authors map the long-term, distributed carbon footprint of the *inference* phase—the ongoing execution of AI queries across global, distributed Application Programming Interfaces (APIs). Furthermore, they account for *embodied* carbon: the emissions generated during the manufacturing, supply chain logistics, and frequent disposal of specialized AI hardware like GPUs.


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

The article exposes a severe fragmentation in environmental disclosure laws, demonstrating how current reporting frameworks allow tech companies to externalize their carbon footprints. By decentralizing inference queries across third-party global APIs, companies exploit regional regulatory disparities. Li et al. argue for a drastic overhaul of corporate transparency mandates, advocating for legally binding, end-to-end carbon accounting that forces AI providers to internalize both the hardware manufacturing emissions and the lifetime query-phase energy expenditures of their models.

**7. Bisoyi, R., General-Purpose AI (GPAI) Obligations and Systemic Risk Assessment Under Article 1 of the EU AI Act, 60 Int'l L. 15 (2025).**

Bisoyi provides a rigorous, doctrinal analysis of the European Union's landmark AI Act, focusing specifically on the statutory obligations imposed on General-Purpose AI (GPAI) models. The article dissects the regulatory mechanisms of Article 1, unpacking how the European AI Office identifies, classifies, and mandates the mitigation of "systemic risks." While systemic risk frameworks traditionally encompass algorithmic bias, mass misinformation, and national security threats, Bisoyi highlights how the EU framework increasingly integrates environmental and infrastructural sustainability into its compliance thresholds.

The core of the analysis evaluates the operational burdens placed on downstream developers and foundation model providers. Under the EU AI Act, providers of models with systemic risks are subjected to stringent transparency mandates, requiring them to document and disclose their energy consumption, compute efficiency, and environmental mitigation strategies. Bisoyi evaluates the extraterritorial "Brussels Effect" of these provisions, arguing that the rigorous compliance standards of Article 1 will functionally dictate international engineering baselines, compelling global tech conglomerates to adapt their software architectures to meet European ecological benchmarks.

**2. The Mechanics of the Paradox: "Green-by-AI" vs. "Red AI"**

The sustainability profile of artificial intelligence is split into two conflicting paradigms:

The environmental impact of artificial intelligence is defined by a deep structural contradiction. To evaluate this "Green AI Paradox" objectively, we must look past the abstract mathematical layer of machine learning and analyze the physical reality of the technology: its energy grids, server nodes, and semiconductor manufacturing pipelines[5] The lifecycle of AI exists in tension between two competing forces: its capacity to optimize planetary systems ("Green-by-AI") and its inherent, resource-heavy computing demands ("Red AI")[6]

**2.1 Green-by-AI: Ecological Contributions**

AI functions as a planetary optimization tool. By processing vast, multi-modal datasets, specialized neural networks achieve high-efficiency outcomes. At the application layer, specialized AI models function as powerful planetary optimization mechanisms. By processing vast streams of multi-modal, unstructured environmental data, these models achieve precise efficiencies that humans cannot match.

Precision Agriculture and Resource Optimization: Traditional farming heavily relies on uniform schedules for watering and chemical distribution, which leads to high resource runoff and nitrous oxide emissions. AI systems process real-time telemetry from Internet of Things (IoT) soil sensors and hyper-spectral satellite arrays to predict precise evaporation rates and soil nitrogen deficiencies [Citation 4: Su, 2026]. This allows automated infrastructure to deliver water and fertilizer only where and when needed, reducing resource waste.

Grid Decarbonization and Renewable Integration: Transitioning to clean energy is often bottle-necked by the unpredictable nature of solar and wind power. Machine learning models resolve this by matching volatile clean generation with real-time consumer demand patterns. Deep learning networks predict grid congestion and storage requirements hours in advance, allowing operators to smoothly integrate clean energy into regional grids without relying on fossil-fuel "peaker" plants.

Conservation Informatics and Automated Remote Sensing: Protecting critical ecosystems requires constant monitoring. Computer vision models are deployed to parse continuous satellite feeds and remote acoustic streams. In rainforests, these neural networks flag the specific frequencies of chainsaws or heavy machinery, alerting local rangers to illegal logging operations days before traditional surveys would detect them[7] .

Precision Agriculture: AI algorithms predict soil moisture evaporation and nitrogen requirements, cutting fertilizer waste and agricultural emissions[8].

Grid Decarbonization: Machine learning maps consumer demand to clean energy output, mitigating the intermittency issues of solar and wind power.

Conservation Informatics: Computer vision models automate the detection of illegal logging and wildlife poaching from satellite and acoustic streams.

[5] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024)

[6] Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025).

[7] Su, X., Green-by-AI: Automated Remote Sensing, Precision Agriculture, and Planetary Optimization Layer Efficiencies, 51 Ecology L.Q. 331 (2026)

[8] Su, X., Green-by-AI: Automated Remote Sensing, Precision Agriculture, and Planetary Optimization Layer Efficiencies, 51 Ecology L.Q. 331 (2026)


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

**2.2 Red AI: The Operational Footprint**

Conversely, the development of generative LLMs is governed by compute-scaling laws where performance increases necessitate exponential expansions of model parameters. This resource-intensive development is known as "Red AI"[9].

The lifecycle emissions of AI are split between embodied carbon—emissions from manufacturing specialized semiconductor hardware like Graphics Processing Units (GPUs)—and operational carbon, generated by the continuous electricity consumed during model training and customer inference queries[10].

While specialized, task-specific models can be relatively lightweight, the rise of generative AI and Large Language Models (LLMs) has shifted the industry toward "Red AI"—a paradigm where algorithmic accuracy is driven by exponential increases in computational power[11]. Modern foundation models are governed by compute-scaling laws: to achieve linear improvements in reasoning capabilities, the underlying training data and model parameters must scale exponentially. This requires vast amounts of physical resources.

[AI Lifecycle Emissions]
├── Embodied Carbon (GPU Manufacturing & Infrastructure Build)
└── Operational Carbon
├── Training Phase (One-time, massive energy spike)
└── Inference Phase (Continuous, token-by-token processing)

While training phase emissions represent an immediate energy spike, the continuous inference phase (answering user prompts) represents the largest long-term carbon footprint due to the sheer volume of global interactions[12].

The true carbon footprint of these systems spans two primary dimensions: embodied carbon and operational carbon[13]

**1. Embodied Carbon and Hardware Infrastructure**

Before an AI model processes a single token of text, it incurs a steep environmental debt through embodied carbon. This includes the emissions generated by mining rare-earth elements, manufacturing advanced application-specific integrated circuits (such as high-end Graphics Processing Units, or GPUs), and constructing massive data centers [Citation 6: Li et al., 2024]. The highly specialized silicon chips required for AI acceleration undergo complex manufacturing pipelines that consume vast amounts of ultra-pure water and energy. Furthermore, the rapid pace of AI hardware innovation creates short depreciation cycles, resulting in high volumes of toxic electronic waste as server clusters are replaced every few years.

**2. Operational Carbon: Training vs. Inference**

Operational carbon represents the emissions generated by the electricity consumed during a model's active lifecycle. This lifecycle is split into two phases with very different energy profiles:

The Training Phase: This phase creates a massive, concentrated surge in energy consumption. Training a foundation model requires thousands of GPUs to run continuously for weeks or months, drawing megawatts of power from local grids [Citation 2: Ebert et al., 2024]. If a server cluster relies on a carbon-heavy grid (such as one dependent on coal or natural gas), the training phase alone can emit hundreds of metric tons of carbon dioxide equivalents (CO_2e). This represents a major, immediate climate impact.

The Inference Phase: While training is a one-time event, inference—the process of an engine answering a live user prompt—is continuous and scales with the user base [Citation 6: Li et al., 2024]. Each individual prompt requires a relatively small amount of energy. However, when scaled across billions of daily global queries, the aggregate energy demand of the inference phase quickly overtakes the total energy spent during training[14].This makes daily usage the dominant driver of AI's long-term carbon footprint.

**3. Data Center Demands: The Cooling and Water Crisis**

Modern high-density AI clusters operate at extreme thermal thresholds, requiring constant cooling to prevent hardware failure. Data centers rely on either energy-intensive air conditioning systems or massive liquid cooling loops. This introduces a dual resource conflict:

Grid Strain: Cooling infrastructure often demands up to 40% of a data center's total energy budget, measured via Power Usage Effectiveness (PUE) [15].

Water Deprivation: When data centers shift to evaporative cooling to save electricity, they consume millions of liters of fresh water daily. This directly strains local water tables and often competes with community water supplies[16].

[9] Tunzina, M. et al., Red AI and Compute-Scaling Laws: The Macro-Level Energy Trajectory of Foundation Models, 2026 Comput. L. Rev. Int'l 45 (2026)
[10] Li, J. et al., Embodied versus Operational Carbon: Mapping Long-Term Inference Phase Emissions Across Distributed Global APIs, 31 Mich. Tech. L. Rev. 201 (2024)
[11] Tunzina, M. et al., Red AI and Compute-Scaling Laws: The Macro-Level Energy Trajectory of Foundation Models, 2026 Comput. L. Rev. Int'l 45 (2026)
[12] Li, J. et al., Embodied versus Operational Carbon: Mapping Long-Term Inference Phase Emissions Across Distributed Global APIs, 31 Mich. Tech. L. Rev. 201 (2024)
[13] Li, J. et al., Embodied versus Operational Carbon: Mapping Long-Term Inference Phase Emissions Across Distributed Global APIs, 31 Mich. Tech. L. Rev. 201 (2024)
[14] Li, J. et al., Embodied versus Operational Carbon: Mapping Long-Term Inference Phase Emissions Across Distributed Global APIs, 31 Mich. Tech. L. Rev. 201 (2024)
[15] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024)
[16] Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

| Parameter / Lifecycle Stage | Core Resource Demand | Environmental Impact |
|---|---|---|
| Model Training Phase | High-density compute over weeks/months | Immediate localized carbon spike |
| Model Inference Phase | Real-time, token-by-token generation | Continuous, scaling global emission footprint |
| Data Center Infrastructure | Liquid/air cooling & constant power | Heavy strain on local grids and water tables [Citation 1: Commins & Irion, 2025] |

**2.2 (Data Centre Demands).**

**The missing argument:** Most papers talk about "energy consumption" broadly, but fail to mention the localized infrastructure bottleneck. Data centers can be built in 12–18 months, but connecting them to a power grid can take 5 to 7 years due to a severe global shortage of high-voltage electrical transformers[17].

This lag has forced tech companies into the "Bring Your Own Power" (BYOP) movement, where they bypass the grid entirely to build private, on-site fossil-fuel or natural gas generation plants[18]. This creates a direct legal and environmental conflict: tech companies are quietly building private fossil-fuel infrastructure behind the meter, completely masking the true energy strain they are placing on national decarbonization goals.

## 3. Legal Contraventions at the International Level

The physical expansion of AI data centers is moving faster than domestic and international environmental law, causing friction with existing treaties.

### 3.1 Undermining the Paris Agreement

The Paris Agreement establishes a legally binding framework to limit global warming to well below 2°C compared to pre-industrial levels. It operates on Nationally Determined Contributions (NDCs), where each nation self-reports and curtails its domestic emissions.

AI development creates severe structural challenges for this regime:

1. **Asymmetrical Grid Strain:** Because AI data centers are highly concentrated—predominantly within specific regions of the United States, Ireland, Germany, and China—they put unexpected strain on localized grids[19]

   [Citation 1: Commins & Irion, 2025]. This sudden surge in baseload electricity demand forces jurisdictions to delay the decommissioning of fossil-fuel plants, directly conflicting with their stated NDCs.
2. **The Carbon Accounting Gap:** Greenhouse gas protocols separate emissions into Scope 1 (direct), Scope 2 (purchased electricity), and Scope 3 (indirect value chain). Many sovereign states lack mandatory frameworks requiring tech firms to isolate and report Scope 3 emissions stemming from outsourced cloud compute clusters[20] [Citation 3: Gupta, 2026]. This renders cross-border AI carbon tracking virtually impossible.

### 3.2 Trans-Border Regulatory Arbitrage

A core legal issue is the emergence of regulatory arbitrage across jurisdictions. AI model development is inherently geography-agnostic; an LLM can be trained on a server cluster anywhere on Earth via simple cloud protocols.

As regions like the European Union enforce stricter energy targets, technology conglomerates are incentivized to route their computational workloads to developing nations or jurisdictions with weak environmental laws and carbon-heavy grids[21]. This allows firms to benefit from lax local enforcement while neutralizing the regulatory efforts of more progressive regions.

## 4. Regional Legislation and Enforcement Gaps

To counter this, specific regional frameworks have attempted to enforce transparency, though they face major limits.

### 4.1 The European Union AI Act and Energy Directives

The European Union's AI Act is the world's most comprehensive direct AI regulation[22] [Citation 2: Ebert et al., 2024]. It treats environmental protection as a core tenet[23] [Citation 7: Bisoyi, 2025]. Under Article 1 of the Act, providers of large, general-purpose AI models are required to document and publicly disclose their model's energy consumption, resource efficiency, and carbon footprint throughout its lifecycle[24] [Citation 7: Bisoyi, 2025; Citation 2: Ebert et al., 2024; Citation 3: Gupta, 2026].

This works alongside the EU Energy Efficiency Directive, which mandates that data center operators report metrics such

[17] Tunzina, M. et al., Red AI and Compute-Scaling Laws: The Macro-Level Energy Trajectory of Foundation Models, 2026 Comput. L. Rev. Int'l 45 (2026)
[18] Gupta, A., Behind the Meter: The Rise of "Bring Your Own Power" and Infrastructure Bottlenecks in Generative AI Data Centers, 15 Columbia J. Eur. L. 89 (2026)
[19] Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025)
[20] Gupta, A., Behind the Meter: The Rise of "Bring Your Own Power" and Infrastructure Bottlenecks in Generative AI Data Centers, 15 Columbia J. Eur. L. 89 (2026)
[21] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024)
[22] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024
[23] Bisoyi, R., General-Purpose AI (GPAI) Obligations and Systemic Risk Assessment Under Article 1 of the EU AI Act, 60 Int'l L. 15 (2025)
[24] Bisoyi, R., General-Purpose AI (GPAI) Obligations and Systemic Risk Assessment Under Article 1 of the EU AI Act, 60 Int'l L. 15 (2025


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

as Power Usage Effectiveness (PUE) and water consumption[25] [Citation 1: Commins & Irion, 2025; Citation 2: Ebert et al., 2024].

The European Union has positioned itself as a pioneer in digital regulation through the formal rollout of the **EU AI Act**, which directly addresses the environmental risks of advanced computation[26] [Citation 2: Ebert et al., 2024; Citation 7: Bisoyi, 2025]. Rather than treating sustainability as a secondary issue, the Act integrates environmental compliance directly into its risk-management frameworks.

**General-Purpose AI (GPAI) Obligations**

The Act establishes strict transparency rules under Chapter V for providers of General-Purpose AI (GPAI) models—foundational systems trained on massive datasets that handle a wide variety of tasks[27] [Citation 3: Gupta, 2026]. The compliance threshold is tied directly to compute capacity: any model trained using computational resources exceeding 10^{25} floating-point operations (FLOPS) is automatically classified as a model with "systemic risk."

Providers of these models must log, verify, and publicly disclose their environmental metrics, including:

- Total energy consumption over the training lifecycle.
- Estimated carbon footprint (CO_2e emissions).
- Model efficiency relative to the hardware infrastructure used.

**Complementary Energy Directives**

The AI Act operates alongside the **EU Energy Efficiency Directive (EED)[28]** [Citation 1: Commins & Irion, 2025; Citation 2: Ebert et al., 2024]. The EED targets the physical layer of computing by requiring data center operators within the EU to report detailed operational metrics to the Common Union Rating Scheme. Facilities must disclose their exact Power Usage Effectiveness (PUE), water consumption effectiveness, and total waste-heat reuse profiles [Citation 1: Commins & Irion, 2025]. This dual approach aims to regulate both the software layer (the models) and the hardware layer (the data centers).

Despite its comprehensive design, the current Western regulatory framework suffers from three major structural gaps that allow tech firms to bypass environmental accountability:

1. The Inference Phase Blind Spot

The most significant limitation of the EU AI Act's sustainability clause is its focus on the initial *training* phase of foundation models[29] [Citation 2: Ebert et al., 2024]. It largely overlooks the ongoing energy consumption of the *inference phase* once a model is deployed to the market. Because inference is distributed across thousands of downstream applications and proprietary APIs worldwide, tracking its real-time carbon intensity remains highly fragmented. By omitting inference from mandatory transparency disclosures, the law ignores the largest source of long-term operational emissions in the AI lifecycle [Citation 2: Ebert et al., 2024; Citation 6: Li et al., 2024].

**2. The Lack of Binding Targets and Rebound Effects**

While current EU directives mandate data collection, they stop short of enforcing hard, statutory caps on fossil-fuel consumption or requiring data centers to procure 100% renewable energy [30]This creates a severe **rebound effect**: as hardware or algorithms become more efficient, the cost of compute drops, which ultimately drives up total demand and net energy consumption [Citation 1: Commins & Irion, 2025]. Without binding energy caps, transparency data simply documents an escalating crisis without legally containing it.

Despite its comprehensive design, the current Western regulatory framework suffers from three major structural gaps that allow tech firms to bypass environmental accountability:

**1. The Inference Phase Blind Spot**

The most significant limitation of the EU AI Act's sustainability clause is its focus on the initial training phase of foundation models [31]It largely overlooks the ongoing energy consumption of the inference phase once a model is deployed to the market. Because inference is distributed across thousands of downstream applications and proprietary APIs worldwide, tracking its real-time carbon intensity remains highly fragmented. By omitting inference from mandatory transparency disclosures, the law ignores the largest source of long-term operational emissions in the AI lifecycle[32]

**2. The Lack of Binding Targets and Rebound Effects**

While current EU directives mandate data collection, they stop short of enforcing hard, statutory caps on fossil-fuel consumption or requiring data centers to procure 100% renewable energy [33]This creates a severe rebound effect: as hardware or algorithms become more efficient, the cost of compute drops, which ultimately drives up total demand and net

[25] Commins, J. & Irion, K., Data Center Demands, Grid Strain, and the Water Crisis: The Physical Cost of Cloud AI, 24 J. Tech. L. & Pol'y 112 (2025)
[26] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024)
[27] Gupta, A., Behind the Meter: The Rise of "Bring Your Own Power" and Infrastructure Bottlenecks in Generative AI Data Centers, 15 Columbia J. Eur. L. 89 (2026)
[28] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024
[29] Ebert, L. et al., The Mechanics of the Green AI Paradox: Evaluating Computational Costs and Regional Regulatory Approaches, 37 Harv. J.L. & Tech. 405 (2024
[30] Commins & Irion, 2025].
[31] Ebert et al., 2024
[32] Ebert et al., 2024
[33] Commins & Irion, 2025


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/

energy consumption[34]. Without binding energy caps, transparency data simply documents an escalating crisis without legally containing it.

### 4.2 Critical Enforcement Gaps

Despite its comprehensive design, the current Western regulatory framework suffers from three major structural gaps that allow tech firms to bypass environmental accountability:

**1. The Inference Phase Blind Spot**

The most significant limitation of the EU AI Act's sustainability clause is its focus on the initial *training* phase of foundation models[35]. It largely overlooks the ongoing energy consumption of the *inference phase* once a model is deployed to the market. Because inference is distributed across thousands of downstream applications and proprietary APIs worldwide, tracking its real-time carbon intensity remains highly fragmented. By omitting inference from mandatory transparency disclosures, the law ignores the largest source of long-term operational emissions in the AI lifecycle[36].

**2. The Lack of Binding Targets and Rebound Effects**

While current EU directives mandate data collection, they stop short of enforcing hard, statutory caps on fossil-fuel consumption or requiring data centers to procure 100% renewable energy[37].This creates a severe **rebound effect**: as hardware or algorithms become more efficient, the cost of compute drops, which ultimately drives up total demand and net energy consumption[38]Without binding energy caps, transparency data simply documents an escalating crisis without legally containing it.

[The Efficiency Rebound Loop]
Algorithmic/Hardware Efficiency Increases
➔ Cost Per Computational Unit Decreases
➔ Market Demand for AI Queries Rises Beneficially
➔ Net Data Center Energy Consumption Escalates
➔ Initial Environmental Gains are Neutralized

**3. Extraterritorial Limitations and Regulatory Safe Havens**

The EU can regulate models placed on its internal market, but enforcing these rules on backend compute operations outside its borders is incredibly difficult. AI model development is highly mobile; code and weight optimizations can be executed on server nodes anywhere on earth via cloud protocols.

If compliance penalties—which can reach up to 15 million Euros or 3% of a company's global annual turnover—become too costly, technology firms can shift their heavy computational work to nations with loose environmental laws and carbon-intensive grids[39]. This cross-border regulatory arbitrage allows companies to bypass strict local laws, undermining global climate goals.

## 5. CONCLUSION AND POLICY RECOMMENDATIONS

The "Green AI Paradox" represents a classic systemic failure: using a carbon-intensive tool to fix a carbon-induced crisis. If left unregulated, the macro-level operational emissions of foundation LLMs risk outstripping the environmental efficiencies gained at the local application layer.

To resolve this imbalance, international governance must transition from voluntary corporate sustainability pledges to binding legal structures:

1. **Unified International Carbon Accounting:** The United Nations Framework Convention on Climate Change (UNFCCC) should establish an explicit digital-infrastructure protocol under the Paris Agreement, forcing nations to integrate state-level AI computational energy consumption into their formal carbon accounting.
2. **Mandatory Inference Tracking:** Expanding on regional laws like the EU AI Act, international policy must mandate that AI platforms provide clear, automated transparency regarding the carbon intensity of individual user queries [Citation 2: Ebert et al., 2024].
3. **The "Green Default" Mandate:** Regulators should compel foundational AI service providers to offer users automated "green routing" settings [Citation 3: Gupta, 2026]. This includes prioritizing execution on server nodes operating on clean grids, using compressed models for simple inquiries, and disabling resource-heavy advanced reasoning when standard processing suffices.

True sustainability cannot survive on algorithmic optimizations alone. It requires global legal frameworks that treat digital infrastructure as a shared planetary resource, ensuring that the advancement of human intelligence does not come at the expense of the global climate.

**Creative Commons (CC) License**

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution–Non-Commercial–No Derivatives 4.0 International (CC BY-NC-ND 4.0) license. This license permits sharing and redistribution of the article in any medium or format for non-commercial purposes only, provided that appropriate credit is given to the original author(s) and source. No modifications, adaptations, or derivative works are permitted under this license.

[34] Commins & Irion, 2025
[35] Ebert et al., 2024
[36] Ebert et al., 2024
[37] Commins & Irion, 2025.
[38] Commins & Irion, 2025
[39] Ebert et al., 2024


© 2026 Dr. Anjana Kumari, Arnav Raviraj, Arya. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND).https://creativecommons.org/licenses/by-nc-nd/4.0/