

Measuring what matters: How to assess AI's environmental impact



Disclaimers

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the International Telecommunication Union (ITU) or of the ITU secretariat concerning the legal status of any country, territory, city, or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by ITU in preference to others of a similar nature that are not mentioned. Errors and omissions excepted; the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by ITU to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader.

The opinions, findings and conclusions expressed in this publication do not necessarily reflect the views of ITU or its membership.

ISBN

978-92-61-40801-5 (PDF version)

978-92-61-40811-4 (EPUB version)



Please consider the environment before printing this report.

© ITU 2025

Some rights reserved. This work is licensed to the public through a Creative Commons Attribution-Non-Commercial-Share Alike 3.0 IGO license (CC BY-NC-SA 3.0 IGO).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited. In any use of this work, there should be no suggestion that ITU endorse any specific organization, products or services. The unauthorized use of the ITU names or logos is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the International Telecommunication Union (ITU). ITU is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition". For more information, please visit <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/>

Measuring what matters: How to assess AI's environmental impact



Foreword by the ITU Deputy Secretary-General



As artificial intelligence (AI) transforms industries, economies, and societies, its rapid growth brings significant environmental implications.

Energy consumption, water usage, and greenhouse gas emissions associated with AI are rising sharply.

Electricity consumption by data centres increased 12 per cent each year from 2017 to 2023, four times faster than global electricity growth, according to

the [IEA](#). As per the [Greening Digital Companies 2025](#) report, four leading AI-focused companies saw their operational emissions increase by 150 per cent on average between 2020 and 2023.

But tech innovation must support, not hinder, climate action.

Digital companies can drive decarbonization. Their technologies offer the potential to boost efficiencies across different industries, unlock innovative climate solutions, and address numerous socio-economic challenges.

To achieve this, the global tech industry needs to better understand and manage positive and negative impacts. The urgency of doing so has never been greater. Yet the potential for greening digital remains obscured by inconsistent measurement practices and fragmented accountability.

That is why ITU and partners worldwide, coming together for [Green Digital Action](#), have called on the tech industry to take responsibility.

With an array of governments, businesses, civil society and international organizations, we aim to establish common, transparent measurement criteria and reporting frameworks, promote and uphold green technology standards, and build a circular economy.

The Declaration on Green Digital Action, put forward by the Azerbaijan COP29 presidency and ITU, has received endorsements from over 80 countries and nearly 1,800 companies and other non-state organizations worldwide. It calls for action to reduce environmental and climate impact while ensuring the benefits of connectivity extend to everyone worldwide.

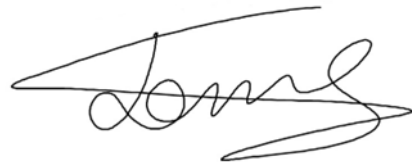
As we approach the next climate talks at COP30, we expect new national and corporate commitments to fully capture digital impact on climate – both negative and positive.

This report, *Measuring what matters: How to assess AI's environmental impact*, reviews common assessment methodologies over the entire AI lifecycle. It stems from the Green Computing pillar of our Green Digital Action initiative, with a key sub-group on sustainable AI.

We are examining implications at each phase of AI development, from initial models and training to real-world usages and interactions.

By identifying gaps and flagging inconsistencies in current AI impact measurement, we aim to provide actionable insights for AI developers, consumers, and policy-makers alike.

Together, let's make sure AI helps us overcome the climate crisis rather than exacerbating it.

A handwritten signature in black ink, appearing to read 'Tomas', with a stylized, sweeping flourish extending from the end.

Tomas Lamanauskas
Deputy Secretary-General
International Telecommunication Union

Geneva, 2025

Foreword by the Director of the ITU Telecommunication Standardization Bureau



International technical standards, developed collaboratively and agreed by consensus, can provide robust frameworks for sustainable artificial intelligence (AI). The findings of this report, however, reveal a field of standardization still in its early stages.

Current approaches to measuring AI's environmental impact across key phases such as development, training, and deployment exhibit significant variability, data gaps, and heavy reliance on estimations.

Indirect impacts from supply chains and hardware manufacturing, for example, are often overlooked or inconsistently assessed. These gaps limit transparency, informed decision-making, and accountability, underscoring the need for more comprehensive and harmonized efforts to align AI development with global sustainability goals.

ITU standards provide metrics and methodologies for climate-impact monitoring. They also help industry integrate emissions tracking in product-development pipelines and offer guidance to governments and companies on lifecycle reporting.

The findings of this report highlight the importance of global action in three areas:

1. **Standardization** of metrics to enable comparisons across studies and stakeholders.
2. **Transparency** on energy, water, and material footprints at every lifecycle stage.
3. **Collaboration** on sustainable AI among developers, consumers, and policy-makers and regulators.

This report assesses the current landscape as a foundation for action. It offers valuable guidance to standards developers and all governments, companies, and researchers committed to integrating sustainability in every phase of AI design and deployment.

A handwritten signature in blue ink that reads "Seizo Onoe". The signature is fluid and cursive, with the first name "Seizo" and the last name "Onoe" clearly distinguishable.

Seizo Onoe
Director, Telecommunication Standardization Bureau
International Telecommunication Union

Geneva, 2025

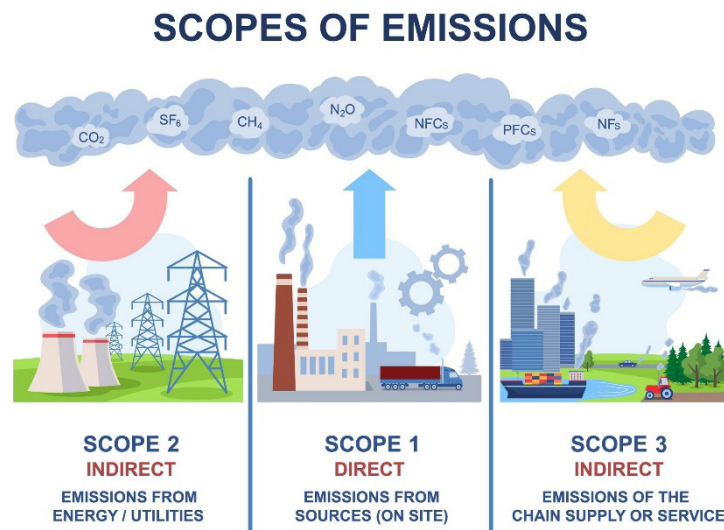
Executive summary

This report synthesizes key findings from a diverse range of sources, including academic literature, corporate sustainability initiatives, and emerging environmental tracking tools. Collectively, these documents provide a thorough overview of current methodologies for evaluating the environmental impacts of artificial intelligence (AI) systems. While several advances in methodology and tooling are evident, the review highlights substantial inconsistencies in how different lifecycle stages of AI are measured, analysed, and reported.

The primary objectives of this review were to assess:

- Which components of AI systems' environmental impact are currently being measured?
- How accurate, transparent, and methodologically sound are these measurement practices?
- How actionable and relevant are these insights for decision-makers, operational teams, and policy-makers?

One of the most pressing issues uncovered is the widespread reliance on indirect estimates when assessing energy consumption during the training phase of AI models. These estimates often lack real-time, empirical measurement. Furthermore, equally important lifecycle stages – such as inference (the operational use of models), Scope 3 emissions (from supply chains and hardware manufacturing), and infrastructure-level impacts (such as water consumption and cooling) – remain significantly underexplored. This reliance on proxies introduces substantial data gaps, impedes accountability, and restricts consumers' ability to make informed, sustainable choices about AI.



Deciphering Carbon Reporting: Exploring Scope 1, 2, and 3 Emissions (image source: [link here](#))

To address these issues, the report uses a lifecycle-based approach, dividing the AI system's environmental impact into three stages:

1. Training,
2. Inference,
3. Supply Chain.

For each stage, we examine measurement methodologies, identify current limitations, and offer recommendations for key stakeholder groups: developers (producers), users (consumers),

and policy-makers. The overarching aim is to ensure that sustainability becomes a foundational element – embedded from the earliest stages of AI design to its deployment and continued use – rather than an afterthought.

Assessing AI's environmental impact

While progress has been made evaluating the environmental impacts of artificial intelligence (AI) systems, significant inconsistencies remain in how different lifecycle stages—such as training, inference, and supply chain—are measured and reported. The review highlights a reliance on indirect estimates rather than direct measurements, especially for energy use during AI training, while other important factors like operational use, supply chain emissions, and infrastructure impacts are often overlooked. By adopting a lifecycle approach and offering targeted recommendations for developers, users, and policy-makers, the report aims to embed sustainability into every stage of AI development and use, ensuring environmental responsibility is prioritized from the outset.

Key AI measurement gaps

Current methods for measuring AI's environmental impact are fragmented and rely too much on estimates instead of real data.

Persistent gaps include:

- **Over-reliance on estimates and proxies**
- **Inconsistent lifecycle boundaries and units**
- **Underreported Scope 3 and embodied emissions**
- **Opaque water use and infrastructure overheads**
- **Neglect of inference phase and user behaviour**
- **Lack of standardization across tools and methodologies**
- **Carbon-centric metrics masking broader impacts**

To further advance the understanding of AI's environmental impact, future research should:

- **Integrate lifecycle assessments**
- **Enhance transparency and accountability**
- **Keep advancing green AI practices**
- **Standardize measurement practices**
- **Improve hardware efficiency**
- **Develop real-time telemetry tools**
- **Model user-behaviour impacts**
- **Refine amortization and attribution methods**
- **Harmonize lifecycle definitions and reporting units**
- **Expand Scope 3 and supply-chain analysis**

Background and purpose

The rapid advancement of AI across enterprise, public, and consumer sectors has raised urgent concerns about its environmental footprint. While high-performance AI systems offer substantial economic and operational benefits, they also generate significant digital emissions due to energy-intensive training, constant inference workloads, and a hardware supply chain dependent on resource-extractive processes.

A key industry challenge is the lack of standardized, transparent metrics to measure AI's environmental impact. This gap results in inconsistent reporting and misleading assumptions and hampers efforts to benchmark and reduce emissions.

Advancing sustainable AI practices

As AI rapidly transforms industries, concerns about its environmental impact—including energy-intensive training, ongoing inference workloads, and resource-heavy supply chains—are growing. The lack of standardized, transparent metrics for measuring AI's footprint leads to inconsistent reporting and impedes progress toward sustainability. Green Digital Action brings together public and private stakeholders to address these challenges. By identifying measurement gaps, developing actionable guidance, and promoting cross-sector collaboration, the initiative aims to provide clear insights into AI's energy use, water consumption, and carbon emissions. This effort empowers organizations to make informed, sustainable decisions about AI deployment, aligning technological progress with environmental stewardship for a greener future.

To address these challenges, the ITU-led Green Digital Action (GDA) initiative and its Green Computing pillar formed a Sustainable AI working group of public and private sector stakeholders. The group aims to identify existing measurement approaches, highlight methodological gaps, and develop actionable guidance for industry-wide adoption to produce quantifiable insights into energy consumption, water usage, and carbon emissions across AI deployment models.

These outputs support strategic decision-making and promote collaboration across sectors to accelerate sustainable AI practices. By exposing misconceptions, the initiative empowers organizations to make informed decisions about AI workload placement and deployment models.

Ultimately, the GDA initiative aims to open the way for a more sustainable future in AI development and deployment, ensuring that technological advancements are aligned with environmental stewardship.



Table of contents

Foreword by the ITU Deputy Secretary-General..... ii

Foreword by the Director of the ITU Telecommunication Standardization Bureau..... iv

Executive summaryv

Background and purpose.....vii

Key findings - AI training..... 1

 i Observed methodologies and best practice 1

 ii Gaps..... 2

 iii Recommendations for developers, consumers, policy-makers/regulators 6

Key findings: AI inference / usage phase..... 7

 i Observed methodologies and best practice 7

 ii Gaps..... 7

 iii Recommendations for developers, consumers, policy-makers/regulators 10

Key findings: AI supply chain and Scope 3 phase 11

 i Observed methodologies and best practice 11

 ii Gaps..... 12

 iii Recommendations for developers, consumers, policy-makers/regulators 12

Gaps in current AI measurement approaches..... 13

Key insights and next steps..... 15

Conclusion..... 17

Project contributors..... 18

Appendix 19

Key findings - AI training

i Observed methodologies and best practice

Energy and carbon footprint of AI training

Training large AI models demands significant computational resources, resulting in high energy use and notable environmental impact. While GPU power draw is often used to estimate energy consumption, accurate assessment must include all system components (CPUs, memory, networking, storage). Best practice is to use actual energy measurements, convert to CO₂ emissions using region-specific grid factors, and consider the model's full lifecycle—including inference—for a complete carbon footprint.

AI training, particularly for large-scale models such as foundation models or transformers, requires substantial computational resources over prolonged periods. This phase represents a significant environmental footprint due to the concentration of energy-intensive hardware operations, typically utilizing GPUs, TPUs, or other accelerators in high-performance computing (HPC) clusters or hyperscale cloud environments.

The most widely observed method for estimating energy consumption during training leverages the following formula:

$$\text{Energy (kWh)} = \text{GPU power draw (kW)} \times \text{Training time (h)} \times \text{Utilization rate}$$

GPU Power Draw is often derived from the Thermal Design Power (TDP), but more accurate measurement is obtained through real-time telemetry using tools that measure actual energy consumption. However, it is critical to note that GPU power draw alone does not represent the full energy consumption of the AI training process. Modern AI training pipelines typically involve additional compute components including CPUs, memory modules, networking devices, and storage systems. These components collectively contribute to the overall energy footprint and are frequently undercounted when only GPU TDP is used.

As a best practice, energy accounting should encompass the total system energy consumption, using observed actuals and not estimates. This energy estimate is then converted to carbon emissions:

$$\text{CO}_2\text{e (kg)} = \sum [\text{total system energy (kWh)} \times \text{Data centre PUE} \times \text{Grid emission factor (kg CO}_2\text{e per kWh)}]$$

Grid emissions factors are highly region-dependent and should ideally be based on marginal rates rather than national averages to more accurately reflect the carbon intensity of consumed electricity at training time. For example, data centres in France may emit far less CO₂e per kWh (~0.05 kg) than those in coal-dependent grids such as Poland (~0.8 kg).

A more granular approach to attributing emissions across a model's lifecycle involves amortization over the model's expected usage volume:

$$\text{Per inference emissions} = \text{Training CO}_2\text{e} / \text{Expected number of inferences}$$

This method is particularly relevant for models deployed as a service, where inference usage can be projected over time. However, assumptions here can significantly affect accuracy, especially for models with variable lifespans or deployment scales.

In addition, commonplace IT infrastructure reporting methodologies are observed as foundational in the reporting of AI environmental impact, for example:

Power usage effectiveness (PUE): Total facility power / IT equipment power

Hyperscale operators (e.g., Google, AWS) report PUEs of 1.1–1.3, compared to traditional data centres averaging 1.6 (Uptime Institute, 2023).

Water usage effectiveness (WUE): Litres of water consumed / IT equipment energy (kWh)

Emerging methodologies aim to benchmark training-phase emissions using standard workloads and hardware comparisons. However, no common approach has yet been observed.



ii Gaps

The methodologies outlined in section 1 do not capture the totality of the training phase environmental impact, and they are not universally applied across the existing literature.

Most of the available tools focus on measuring electricity usage during the inference phase of AI models, typically expressing emissions in units of mass of carbon per request, omitting the training phase altogether. Model training has been included in several lifetime carbon assessment (LCA) studies (e.g. Berthelot et al 2024; de Vries, 2023), although there is variation in precisely what was measured, and even in how the training phase is defined across the studies we reviewed.

Key measurement gaps for AI environmental impact

Current methodologies fail to fully capture AI training's environmental footprint due to:

1. Over-reliance on proxies and estimates – Energy/emissions data often uses estimates instead of real-time telemetry, creating high uncertainty.
2. Inconsistent boundaries – Definitions of "training" vary significantly, frequently excluding R&D phases with substantial cumulative emissions.
3. Unstandardized units – Metrics range from emissions per token to per training cycle, hindering comparability and cost allocation.
4. Carbon myopia – Assessments overwhelmingly focus on CO₂ emissions, neglecting water use, e-waste, mineral depletion, and biodiversity impacts.

We identify the following barriers to measuring the environmental impacts of model training:

Over-reliance on proxies and estimates

Many studies report training emissions calculated from estimated values for model runtime, GPU power draw, carbon intensity factors, rather than real telemetry or direct measurement.

There is a scarcity of real data disclosed by model producers and data centre operators, meaning impact assessors fall back to proxy measurements and estimates for key values including, but not limited to, details of the hardware used, training time, amount of energy consumed and its carbon intensity.

There is significant uncertainty in these estimates, which propagates through to high uncertainty in the final measurements. The over-reliance on proxies and estimates of unknown accuracy is especially acute for the embodied impacts of the training hardware, where it is typical to apply generic models to assumed hardware configurations.

Inconsistent "training" boundary

Training emissions are not always included in environmental assessments, despite their importance. Furthermore, the definition of the training phase is inconsistent across the literature.

For example, the research and development phase, which often involves multiple rounds of model training, refining, and ablation, is usually omitted—even though its cumulative emissions can exceed those of the final training run.

Only a handful of studies attempt to break down emissions using the phases of the AI development lifecycle (Build-Train-Run). The absence of phase-based modelling prevents teams from understanding where emissions are concentrated and from optimizing design-stage decisions.

Unstandardized units

Estimates of the environmental impacts of AI models express their results in a range of units that are not easily comparable or interchangeable. Different metrics are appropriate for different audiences. For example, units related to the model function, such as:

- emissions per token
- emissions per inference
- emissions per task completed
- emissions per user session

are useful for developers or operations teams that aim to reduce their impacts by changing how they interact with a model.

On the other hand, units such as:

- emissions per unit time
- emissions per deployed instance
- emissions per training cycle
- emissions per user

are useful for model providers concerned with managing costs or regulatory compliance.

This problem is especially acute in organisations with internal carbon pricing or cost allocation models, where emissions need to be traceable to the workload, team, or business function responsible.

Without allocatable metrics, emissions remain abstract. They cannot be embedded into procurement, architecture planning, or continuous improvement processes.

Expressing the carbon emitted during the training phase in units other than total carbon footprint is especially challenging because there is uncertainty around how to properly amortize the training carbon so that it can sum with carbon from the other life cycle phases.

Carbon myopia: Overemphasis on CO₂

Where training is accounted for in impact assessments, typically only carbon emissions are included, with other impacts such as energy consumption, water use, biodiversity loss, e-waste and mineral extraction only occasionally included. LCA-based studies usually express environmental impacts across three dimensions: global warming potential (GWP) in units of kgCO₂e (where CO₂e is carbon dioxide equivalent), abiotic depletion potential (ADP) in the unit kgSb_e (kilograms of antimony-equivalent, with antimony being a chemical element used to represent raw material consumption), and primary energy (PE) in megajoule (MJ). By converting to carbon emissions, results can be hidden in carbon accounting methodologies e.g. offsetting with market-based emissions. Instead, the recommendation is to provide the base units of the emission sources such as energy consumption, water usage, minerals consumed etc, so that the carbon emission working can be shown from the base sources.

Positive examples

The estimation of the environmental impact of Generative AI conducted by de Vries (2023), Berthelot et al (2024) and Huang et al (2025) offers guidance for measuring AI model emissions that include the training phase. However, these examples all vary in their implementation details and there is no clear way to choose between them. Typically, training emissions are likely to be most important for attributional studies, LCA studies, and periodic and ongoing reporting, rather than point in time assessments aim to empower consumers to make sustainable choices about their AI usage, because the training emissions historical and cannot be influenced by present and future behavioural changes, whereas inference emissions can.

Recommendations for measurement

Establishing lifecycle-based frameworks and clear boundaries for what constitutes the training phase is recommended for more accurate and actionable reporting. More detailed disclosures from model producers and data centres about the energy, hardware and time allocated to model training would greatly benefit environmental impact assessments. In the absence of that data, guidance for appropriate proxies or heuristics that can be used will help to standardize the measurement landscape.

Gap	Implications	Recommendations
Lack of standardization for incorporating R&D and experimentation cycles into total emissions	Underestimates total training impact; distorts carbon accounting across organizations	Develop a unified framework to include all training iterations and tuning steps in energy reporting, including transfer learning from previous models
Omission of infrastructure overheads (e.g., PUE adjustments)	Incomplete lifecycle assessment: energy use appears artificially low	Include Power Usage Effectiveness (PUE) factors and infrastructure telemetry in all reporting
Reliance on estimates rather than measured telemetry	Results in potentially inaccurate energy and emissions figures	Use actual consumption figures, avoid the use of proxies, estimates and global averages
Absence of standardized amortization models for per-inference attribution	Hinders comparability across services and models	Establish sector-wide amortization guidelines tied to expected usage and model lifespan
Inconsistent units	Difficult to compare different studies	Develop consensus around standard units for specific use cases and guidance for converting between them
Inconsistent life cycle phase definition	Hard to compare studies with different phase definitions	Establish standard definitions for life cycle phases

iii Recommendations for developers, consumers, policy-makers/regulators

Stakeholder	Recommendation	Suggested Action	Next Steps
Developers	Implement real-time emissions tracking tools in development pipelines and standardise reporting and benchmarking with each model release	During training workflows log and report actual energy usage	Integrate emissions monitoring in software CI/CD pipelines and develop model incorporating lifecycle emissions Standardise and benchmark results with each new model release
Consumers	Demand full-lifecycle carbon disclosures for model procurement	Require emissions data from vendors before onboarding models	Incorporate sustainability criteria in procurement and vendor assessment processes, report internally on observed emissions resulting from model usage
Regulators	Define standard emissions accounting procedures, including amortization and infrastructure inclusion	Issue guidelines aligned with ISO 14040, ITU-T L.1410 and CSRD for AI lifecycle reporting	Launch consultations with industry stakeholders to co-develop mandatory reporting standards

Key findings: AI inference / usage phase

i Observed methodologies and best practice

Inference, or the operational deployment of AI models, is often less energy-intensive per unit activity than training but may contribute more to total emissions due to high frequency and long-term use. Measurement at this stage is essential for service-oriented AI products, such as LLM-powered APIs or recommendation engines.

The foundational formula for evaluating inference energy usage is:

$$\text{Energy per inference (Wh)} = \text{Total runtime energy (Wh)} / \text{Inference count}$$

For tokenized models (e.g., large language models), a more refined metric is:

$$\text{Energy per token (Wh)} = (\text{Power} \times \text{Inference duration}) / (\text{Number of tokens} \times 3600)$$

Such metrics enable per-output tracking, which is increasingly useful for quantifying the marginal cost of generative AI tasks. Telemetry tools like can be embedded into model-serving infrastructure to track real-time usage.

Advanced metrics observed include:

Per query emission allocation

$$\text{CO}_2\text{e/query} = (\text{Inference energy} + \text{Overhead energy}) \times \text{local grid emission factor} / \text{Queries processed}$$

Energy performance efficacy

$$\text{Energy efficiency score} = \text{Model accuracy} / \text{Energy consumed}$$

This is valuable when comparing hardware or model architecture choices.

ii Gaps

The methodologies outlined in Section 1 only cover the electricity consumed by an AI model serving inference requests. There are several associated limitations to the methods and how they are applied in AI model impact assessments:

Lack of real-world measurement

Few sources attempt to measure or model emissions during real-world inference. Most rely on estimates or proxies rather than actual telemetry data. There is a widespread use of assumed model runtime, generic hardware, and standard carbon intensity factors, which introduces significant uncertainty.

Platform and deployment assumptions

Many papers assume that cloud deployment is automatically more sustainable, but this is rarely tested or broken down by region, hardware, or workload efficiency. The impact of deployment choices (e.g., cloud region, hardware type, workload placement) on inference emissions is not

well studied. Embodied carbon is rarely available for the specific hardware being used, leading assessors to fall back to analogues or crude models or omit it altogether.

User behaviour and consumption patterns

User behaviour (e.g., prompt length, retries, unnecessary use of large models) is a major factor in overall inference emissions but is almost completely absent from current analyses. There is no exploration of how product design or usage controls could reduce inference emissions.

Lifecycle phase definition

Inference typically considers the electricity consumed server-side to generate a response, and sometimes the hardware used to serve inference. However, inference can also include end-user devices, end-user energy consumption and the operational and embodied carbon emitted to transfer information between the client and server.

Overemphasis on carbon emissions

Where training is accounted for in impact assessments, typically only carbon emissions are included, with other impacts such as water use, biodiversity loss, e-waste and mineral extraction only occasionally included. LCA-based studies usually express environmental impacts across three dimensions: global warming potential (GWP) in units of kgCO_2e (where CO_2e is carbon dioxide equivalent), abiotic depletion potential (ADP) in kgSb_e (kg antimony-equivalent), and primary energy (PE) in MJ.

Inconsistent units

Similarly to the training phase, estimates of the environmental impacts of AI model during the inference phase express their results in a range of units that are not easily comparable or interchangeable.

Positive examples

The EcoLogits Calculator integrates with real development workflows via CodeCarbon. It provides developers with near-real-time visibility into the emissions impact of their workloads during development or model experimentation. The AI Energy Score (Huggingface/Salesforce) runs models on benchmark hardware and measures their energy consumption, adding them to a leaderboard.

These examples indicate what good could look like:

- Feedback loops embedded in coding and training environments
- Configurable granularity by workload, developer, or session
- Integration with Jupyter Notebooks, GitHub Actions, or ML platforms

These types of tools represent the bridge between broad emissions estimates and actionable insights. However, they remain isolated examples rather than widely adopted best practice, and only a minority of models have been measured using these tools, meaning the majority of AI energy consumption remains a blind spot. These tools also only assess energy consumption.

Recommendations

Develop real-time telemetry and integration of environmental measurements into workflows, including inference. Create allocatable metrics (e.g., per-inference, per-token) to make emissions data actionable for teams, and comparable with cost data. Benchmark deployment options to understand the impact of different clouds, regions, and architectures on inference emissions. Encourage responsible design and user behaviour to optimize inference efficiency.

Gap	Implications	Recommendations
Minimal telemetry on inference load	Prevents accurate attribution of operational carbon emissions	Deploy real-time energy monitoring tools in serving infrastructure
User behaviour (prompt length, retries) not modelled	Emissions intensity can vary significantly based on usage patterns. Lack of emissions feedback perpetuates inefficient usage patterns.	Develop behavioural usage models for more granular emissions reporting
No standardized benchmarks for inference across providers or regions	Difficult to compare service carbon footprints	Establish industry-wide inference benchmarking frameworks
Regulatory void	No mandates for inference-phase disclosures in ESG reports	Consider expanding and aligning CSRD's Digital Emissions Reporting standards
Inconsistent units	Difficult to compare different studies	Develop consensus around standard units for specific use cases and guidance for converting between them
Inconsistent life cycle phase definition	Hard to compare studies with different phase definitions	Establish standard definitions for life cycle phases

iii Recommendations for developers, consumers, policy-makers/regulators

Stakeholder	Recommendation	Suggested Action	Next Steps
Developers	Embed telemetry in model-serving APIs	Log inference emissions per request	Add environmental observability to monitoring dashboards and APIs so that consumers can report on their environmental impact Provide guidance on best practice for the consumers usage of AI models, so that they can reduce their environmental footprint
Consumers	Use emissions-aware interfaces to guide AI usage	Provide feedback on high-emissions queries	Integrate emissions into user-facing analytics tools
Regulators	Develop standards for usage-phase emissions reporting	Require disclosure of CO ₂ e/token or CO ₂ e/request for public-facing AI tools	Coordinate with cloud service providers and ML vendors to harmonize metrics

Key findings: AI supply chain and Scope 3 phase

i Observed methodologies and best practice

Scope 3 emissions represent upstream and downstream lifecycle impacts, including manufacturing, transportation, and disposal of AI infrastructure. Estimating these impacts requires integration of product lifecycle assessments (including lifetime carbon assessments, or LCAs), as well as an understanding of how life cycle stages affect Scope 3 emissions calculations and reporting.

The primary equation used is:

$$\text{Embodied carbon (CO}_2\text{e)} = \sum (\text{Mass of components} \times \text{Emission factors})$$

Circular economy metrics:

$$\text{Recycling rate (\%)} = \sum (\text{Recycled materials} / \text{Total materials}) \times 100$$

Spend based accounting

$$\text{Scope 3 emissions} = \sum (\text{Procurement spend} \times \text{Emission factor sector})$$

ISO 14040/14044 provides a structured framework for conducting cradle-to-grave, LCA-based lifecycle assessments. All material & energy inputs, and output products such as waste and emissions are accounted for across all processes and all stages of the AI system lifecycle. However, in observed AI sustainability reports, the emissions from manufacturing and the broader supply chain are regularly missing.



ii Gaps

Gap	Implications	Recommendations
Lack of transparency in hardware sourcing and manufacturing	Makes embodied carbon and social impacts invisible	Mandate supplier disclosures, reporting and third-party certifications
Inconsistent tracking of water use and e-waste	Misses key environmental factors, especially in water-stressed regions	Report WUE and hardware recycling metrics with operational data
Fragmented standards for Scope 3 assessment	Prevents comparability across vendors and deployments	Harmonize LCA protocols using ISO and ICT-specific extensions see also ITU-T L.1410
Data scarcity	Inconsistent reporting on mining impacts (e.g. cobalt, lithium etc)	Develop and share open datasets

iii Recommendations for developers, consumers, policy-makers/regulators

Stakeholder	Recommendation	Suggested Action	Next Steps
Developers	Adopt ISO-compliant LCA tools for infrastructure	Use LCA data in supply chain impact models	Build a component-level environmental product declaration (EPD) system and integrate into an ICT sustainability database
Consumers	Procure AI services with verified low Scope 3 impacts	Select vendors with circular economy certifications	Include Scope 3 in sustainability procurement criteria
Regulators	Extend reporting standards to include full lifecycle and embodied impacts	Add Scope 3 AI-specific disclosures to CSRD or SEC ESG reporting	Develop ICT-specific addenda to global sustainability standards

Gaps in current AI measurement approaches

Key gaps in AI environmental measurement

Current methods for measuring AI's environmental impact are fragmented and rely too much on estimates instead of real data. Important impacts such as supply chain emissions, water use, and hardware lifecycles are often missed. Reporting standards and units vary, making comparisons hard. Most assessments focus narrowly on carbon, ignoring broader environmental effects. Standardization and broader metrics are urgently needed for true accountability.

The current methodologies for assessing the environmental impact of AI reveal important gaps that impede a holistic understanding of their environmental footprint. Despite improvements in measuring energy consumption, the translation of these into broader environmental metrics remains inconsistent. The complexity of AI programs, coupled with their rapid evolution, necessitates a more holistic approach to measurement that encompasses the full AI lifecycle and diverse environmental impacts.

Current methodologies for assessing AI's environmental impact remain fragmented, inconsistent, and incomplete. Despite emerging tools and growing awareness, key lifecycle stages and impact dimensions are underrepresented or poorly quantified.

Persistent gaps include:

Over-reliance on estimates and proxies

Many assessments use forecasted values rather than measured telemetry, introducing uncertainty. Hardware details, energy use, and carbon intensity are often inferred, not observed.

Inconsistent lifecycle boundaries and units

Definitions of lifecycle phases (e.g., "training" vs. "development") vary, complicating comparison. Impact units differ across studies (e.g., per-token vs. per-inference vs. per-session), limiting interoperability and practical use.

Underreported Scope 3 and embodied emissions

Supply chain impacts—such as emissions from chip fabrication, hardware transport, and e-waste—are often excluded. Embodied carbon and mining impacts are poorly tracked and inconsistently reported.

Opaque water use and infrastructure overheads

Water usage data remains sparse and non-standardized, especially from hyperscale providers. Energy overheads like cooling (PUE) are frequently omitted from model-level assessments.

Neglect of inference phase and user behaviour

Inference emissions are often overlooked despite their cumulative scale. Usage patterns, prompt retries, and inefficient deployment choices (e.g., region, hardware) are rarely factored in.

Lack of standardization across tools and methodologies

Disparate tools and reporting formats hinder comparability. There is no consensus on amortization methods, benchmarking protocols, or lifecycle accounting models.

Carbon-centric metrics masking broader impacts

Most studies focus narrowly on CO₂e, ignoring water, minerals, biodiversity loss, and energy source transparency. Results may be obscured by offsets or averaged emissions factors.

Key insights and next steps

The analysis of existing literature, measurement approaches, and identified gaps provides valuable insights into the current state of AI's environmental impact assessment and suggests priority areas for further research and improvement.

Insights from recent studies

- **Energy and resource metrics:** Current research emphasizes the importance of measuring energy consumption in megawatt-hours (MWh) and terawatt-hours (TWh), as well as water usage in litres per kilowatt-hour (L/kWh). However, there is a need for more comprehensive carbon footprint data that translates these metrics into CO₂-equivalent emissions.
- **Scope 3 emissions:** While acknowledged, Scope 3 emissions related to supply chain impacts remain underreported. This gap underscores the need for more empirical data to fully capture the lifecycle environmental impact of AI technologies.
- **Methodological transparency:** The reliance on actual corporate sustainability reports enhances credibility, yet future projections and water impact estimations require greater methodological transparency to ensure accuracy and comparability.

Focus areas for further studies

To address the identified gaps and enhance measurement practices, the next phase of the study proposes to focus on identifying key AI use cases across industries and selecting deployment scenarios to gain a better understanding of diverse environmental impacts. This will include developing a comprehensive framework for testing and measuring and monitoring energy consumption, water usage, and carbon emissions of AI workloads, while also defining variables for workload placement, hardware selection, and operational practices.

In terms of measurement and analysis, the study recommends assessing anonymized operational data related to AI workloads to evaluate environmental impacts across the full lifecycle, and conduct comparative analyses across on-premises, cloud, and hybrid environments to identify best practices and areas for improvement.

Ongoing standardization work

ITU within ITU-T Study Group 5 "Environment, EMF, Climate Action and Circular Economy" is working on a methodology on how to assess the Environmental Impact of Artificial Intelligence systems.

This standard will include:

- Overview of the impacts of AI on environment
- Solutions and framework for evaluating the environmental impact of AI
- Guidance for comparative analysis between two AI systems or between an AI system and a non-AI system.
- Considerations on environmental scoring.

Recommendations for future research

To further advance the understanding of AI's environmental impact, future research should:

- **Integrate lifecycle assessments:** Adopt a life-cycle-based system thinking approach to assess and design AI systems with environmental sustainability in mind
- **Enhance transparency and accountability:** Introduce specific environmental disclosure requirements to enhance transparency across the AI supply chain and ensure comprehensive reporting of energy use and emissions.
- **Keep advancing green AI practices:** Encourage research into methods for greening AI and promote widespread adoption of green AI principles, leveraging AI to address its own environmental challenges.
- **Standardize measurement practices:** Improve AI energy measurement standardization, advocating for direct energy measurements over proxy estimates and ensuring comprehensive power tracking across computing infrastructure.
- **Improve hardware efficiency:** Use more energy-efficient AI hardware to reduce emissions per computation and ensure reporting is transparent and without proxies.
- **Develop real-time telemetry tools:** Advance research into telemetry-based tools that provide fine-grained, real-time insights into emissions at the workload, session, or model level.
- **Model user-behaviour impacts:** Study how user interaction patterns (e.g., prompt length, retry frequency) affect AI energy consumption and explore how interface design can encourage low-emission usage.
- **Refine amortization and attribution methods:** Investigate standard models for amortizing training emissions and attributing environmental costs across services, use cases, and lifecycle phases.
- **Harmonize lifecycle definitions and reporting units:** Establish consensus on lifecycle phase boundaries (e.g., Build-Train-Run) and recommended units for impact reporting (e.g., per-token, per-session, per-task).
- **Expand Scope 3 and supply-chain analysis:** Support detailed studies of embodied emissions, water usage, mineral extraction, and end-of-life impacts associated with AI hardware, including cloud infrastructure and end-user devices.

Conclusion

In conclusion, this paper underscores the critical importance of addressing the identified gaps in measuring the environmental impact of AI. By enhancing transparency, standardizing measurement practices, and integrating lifecycle assessments, the industry can gain a more comprehensive understanding of AI's environmental footprint.

Improved measurement practices will not only enable informed decision-making for optimizing AI operations but also drive the adoption of sustainable AI practices across various sectors. Ultimately, these efforts will contribute to a more sustainable future in AI development and deployment, aligning technological advancements with environmental stewardship.

Project contributors

ITU's Green Digital Action team acknowledges the valuable expertise and insights of the following contributors to this publication:

Name	Organization	Contribution
Thomas Basikolo	ITU	Co-lead
Tim Smolcic	HSBC	Co-lead
Nathalie Otte	Decathlon	Drafting
Joseph Cook	Green Software Foundation	Drafting
Jerry Huang	Harvard University	Drafting
Nadim Kapadia	HSBC	Drafting
Karl Ni	MasterCard	Drafting
Nataliya Tkachenko	Lloyds Banking Group	Drafting
Mark Butcher	Posetiv	Drafting
Mohan Gandhi	Posetiv	Drafting
Blessed Guda	ITU	Review
Joe Murray	Ernst & Young	Review
Kevin Thompson	Global Enabling Sustainability Initiative (GESI)	Review
Jon Turnbull	Google	Review
John Ridd	Green Pixie	Review
Tony Darcy	Nokia	Review
Reyna Ubeda	ITU	Review
Neil MacDonald	ITU	Review
Robin Zuercher	ITU	Review

Appendix

Studies and publications relating to the measurement of AI's environmental impact

No	Content piece name/title	Link to source
1	A Beginner's Guide to Power and Energy Measurement and Estimation for Computing and Machine Learning	https://urldefense.com/v3/https://www.nrel.gov/docs/fy25osti/91518.pdf;!!LSAcJDIP!0MR0U1GXxhVub78Hb-GYsL97y9GTauGMCPTe7XHzA7afhzrl9NBfdf4MxqRM23DpijUuRtKVuqG4JDQ\$
2	A Dataset for Research on Water Sustainability	https://dl.acm.org/doi/10.1145/3632775.3661962
3	AFNOR SPEC 2314 Frugal AI referential	https://www.boutique.afnor.org/en-gb/standard/afnor-spec-2314/fa208976/421140
4	AI threats to climate change	https://foe.org/wp-content/uploads/2024/03/AI_Climate_Disinfo_v6_031224.pdf
5	Artificial intelligence, data, calculations: what infrastructures in a decarbonized world?	https://theshiftproject.org/wp-content/uploads/2025/03/2025_03_06-TSP-Rapport-intermediaire-IA.pdf
6	Artificial intelligence, data, calculations: Key figures (World)	https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Ftheshiftproject.org%2Fwp-content%2Fuploads%2F2025%2F03%2FVERSION-PROVISOIRE-Etat-des-lieux-Monde-The-Shift-Project-1.xlsx&wdOrigin=BROWSELINK
7	A Water Efficiency Dataset for African Data Centers	https://arxiv.org/abs/2412.03716
8	Digital Economy report 2024 - UN	unctad.org/system/files/official-document/der2024_en.pdf
9	Estimating the environmental impact of Generative-AI services using an LCA-based methodology	https://www.sciencedirect.com/science/article/pii/S2212827124001173?ref=pdf_download&fr=RR-2&rr=94750b04f9f603b1
10	EcoLogits Calculator	https://huggingface.co/spaces/genai-impact/ecologits-calculator
11	Fine-grained methodology to assess environmental impact of a set of digital services	https://hal.science/hal-04928998
12	IA Act Regulation Europe	https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206
13	Improving Carbon Emissions of Federated Large Language Model Inference through Classification of Task-Specificity	https://hotcarbon.org/assets/2024/pdf/hotcarbon24-final109.pdf
14	ISO normalization ISO/IEC 42001:2023	https://www.iso.org/fr/standard/81230.html
15	ITU AI & Environment report	ITU AI and Environment Report.pdf

(continued)

No	Content piece name/title	Link to source
16	Life-Cycle Emissions of AI Hardware: A Cradle-To-Grave Approach and Generational Trends	Life-Cycle Emissions of AI Hardware: A Cradle-To-Grave Approach and Generational Trends
17	Measuring trends in Artificial Intelligence	https://aiindex.stanford.edu/ai-index-report-2023/
18	Mineral Resources impact of digital acceleration (AI impact)	Metal requirements in the digital sector - La librairie ADEME
19	Open LLM Leaderboard	https://huggingface.co/spaces/open-llm-leaderboard/open_llm_leaderboard#/
20	Recommendation ITU-T L.1410: Methodology for environmental life cycle assessments of information and communication technology goods, networks and services	https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=16010
21	Recommendation ITU-T L.1480: Enabling the Net Zero transition: Assessing how the use of information and communication technology solutions impact greenhouse gas emissions of other sectors	https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=15030
22	RIA31 Ethical and Responsible AI Guide	https://ref-ia.isit-europe.org/#
23	Standardization for AI Environmental Sustainability – Towards a coordinated global approach	https://www.sustainableaicoalition.org/wp-content/uploads/Standardization_AI_Sustainability.pdf
24	SBTI methodology for IT	https://sciencebasedtargets.org/resources/legacy/2020/04/GSMA_IP_SBT-report_WEB-SINGLE.pdf
25	Towards Green AI: Current Status and Future Research	[2407.10237] Towards Green AI: Current status and future research
26	The Environmental Impact of AI	https://www.splunk.com/en_us/blog/learn/ai-environmental-impact.html
27	The growing energy footprint of artificial intelligence	https://www.cell.com/joule/fulltext/S2542-4351(23)00365-3
28	The growing energy footprint of artificial intelligence	https://www.cell.com/joule/fulltext/S2542-4351(23)00365-3
29	The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink	https://ieeexplore.ieee.org/document/9810097
30	Website Carbon Impact	https://greencompute.uk/References/Web/Website_CarbonImpact_Shina.pdf

(continued)

No	Content piece name/title	Link to source
31	White Paper on Global Artificial Intelligence Environmental Impact (Green AI Index: A Framework for Evaluating the Environmental Impact of AI Models and Data Centres)	https://www.greenai.institute/whitepaper/white-paper-on-global-artificial-intelligence-environmental-impact

International Telecommunication Union
Place des Nations
CH-1211 Geneva 20
Switzerland

ISBN 978-92-61-40801-5



Published in Switzerland
Geneva, 2025

Photo credits: Adobe Stock