Evaluating Articial Intelligence Beyond Performance

Supplement #1: Ethical AI & Sustainable Computing: A Comparative Analysis of Transparency and Accountability Among Leading AI Entities - April 2025

By TowerIO LLC

Evaluating Articial Intelligence Beyond Performance

Forward

The rapid evolution of Artificial Intelligence (AI) presents both unprecedented opportunities and profound challenges. To navigate this complex landscape responsibly, we require robust evaluation frameworks that extend beyond traditional performance metrics. This report arrives as a necessary sequel and supplement to our previous work, building upon its foundation to address a widening range of critical concerns.

Our initial exploration revealed a systemic deficiency in the tools and methodologies available for comprehensive AI assessment. The overwhelming emphasis on metrics like speed, accuracy, and output quality often overshadows equally vital dimensions: ethical considerations, encompassing fairness, transparency, and accountability; environmental sustainability, with its urgent focus on energy consumption and resource stewardship; and the mission-critical imperatives of accessibility, equity, and responsible governance. This report expands upon that initial investigation, delving deeper into these multifaceted challenges and advocating for a more holistic and rigorous approach.

The methodologies herein are not merely academic exercises. They are essential tools for navigating a technological landscape that demands careful, informed decision-making. As AI systems become more deeply integrated into our lives, influencing everything from healthcare and education to employment and governance, the potential for both benefit and harm grows exponentially. We can no longer afford to evaluate these systems solely through a narrow, technical lens. We must adopt a comprehensive framework that accounts for the complex interplay of ethical, social, and environmental factors.

This report synthesizes insights from diverse fields, forging a transdisciplinary approach to AI evaluation. It seeks to provide researchers, policymakers, developers, and all stakeholders with the knowledge and tools necessary to demand and enact greater accountability from those who create and deploy AI technologies.

Executive Summary

Purpose: This report provides a comparative analysis and ranking of 25 key entities operating within the artificial intelligence (AI) ecosystem. The assessment focuses on their publicly demonstrable ethical standards and environmental sustainability practices, with a primary emphasis on transparency and accountability, based on information available as of [Current Date]. The entities analyzed represent a cross-section of the AI landscape, including major technology corporations, specialized hardware manufacturers, AI-focused startups, cloud service providers, and prominent research institutions.

Key Challenge: A fundamental challenge underpinning this analysis is the significant lack of standardized, comparable, and independently verified data across the AI industry. Information regarding specific ethical performance metrics (such as the outcomes of internal bias audits) and the precise environmental footprint of AI operations (like energy or water consumption per AI task) is often proprietary, inconsistently reported, or aggregated within broader corporate disclosures. Consequently, this assessment relies heavily on publicly available information, including company reports, official statements, policy documents, responses to documented incidents, and third-party analyses, using transparency and accountability as primary proxies for underlying practices.

Methodology Overview: A structured assessment framework employing a multi-dimensional scoring rubric was developed and applied to each of the 25 target entities. Dimensions evaluated include: Ethical Governance Transparency & Consistency, Ethical Accountability & Responsiveness, Bias & Fairness Transparency, Environmental Reporting Specificity & Verification (focused on AI), Environmental Action & Commitment (focused on AI), and Overall Transparency Willingness. Scores were assigned based on the availability, clarity, and substance of public evidence for each dimension.

Core Findings Highlights: The analysis reveals significant variance in transparency and accountability across the AI sector. Large, publicly traded technology companies (e.g., Google/Alphabet, Microsoft, IBM) and certain well-funded AI labs (e.g., Anthropic, OpenAI) tend to provide more extensive public documentation regarding AI principles, governance structures, and, to a lesser extent, responses to ethical challenges. However, even among these leaders, a gap often exists between stated principles and documented practices or incident outcomes. AI-specific environmental reporting remains a major area of opacity across the board, with most companies aggregating AI impacts within overall corporate sustainability data. While hardware

efficiency gains are widely promoted, the absolute environmental footprint of AI continues to grow, driven by data center expansion and increasing Scope 3 emissions, particularly from hardware manufacturing and construction, which are often inadequately addressed in public commitments. Research institutions generally demonstrate openness in their research principles but provide less transparency regarding their operational environmental impact.

Critical Limitations: The findings and rankings presented herein must be interpreted with significant caution due to inherent limitations. The analysis is constrained by its reliance on publicly available, often self-reported, information, much of which lacks independent verification. Inconsistent metrics and reporting methodologies hinder direct comparability between entities. Furthermore, the rankings may reflect a company's willingness and capacity to communicate and report as much as, or potentially more than, their actual ethical and environmental performance. The AI field is highly dynamic, and this report represents a snapshot at a specific point in time; company practices and disclosures are subject to rapid change. Greater mandatory, standardized, and independently verified reporting is essential for more definitive future assessments.

I. Introduction: The AI Transparency & Sustainability Imperative

The rapid development and deployment of artificial intelligence represent one of the most transformative technological shifts of the modern era. AI offers profound potential benefits across nearly every sector, from accelerating scientific discovery and enhancing healthcare to optimizing industrial processes and personalizing user experiences. However, this transformative power is accompanied by significant ethical considerations and substantial environmental costs. Issues such as algorithmic bias leading to discrimination, threats to privacy through data collection and analysis, concerns over the safety and reliability of autonomous systems, questions of accountability when AI systems cause harm, and the potential for misuse demand careful management.[1] Simultaneously, the computational intensity of training and deploying sophisticated AI models, particularly large language models (LLMs), drives considerable energy consumption, water usage for cooling data centers, and contributes to electronic waste, placing additional strain on planetary resources.[4]

In response to these dual challenges, the concept of "Responsible AI" has emerged as a critical framework. Responsible AI encompasses the ethical design, development, deployment, and governance of AI systems, explicitly including considerations for environmental sustainability.[29] A growing chorus of stakeholders—including the public, policymakers, investors, and employees—is demanding greater transparency and accountability from organizations developing and deploying AI technologies.[42] They seek assurance that AI is being created and used in ways that align with societal values, respect human rights, and minimize environmental harm.

However, assessing the actual ethical performance and environmental sustainability of AI actors presents a significant challenge. As outlined in the user query, there is a fundamental lack of standardized, comparable, and, crucially, independently verified data across the industry [Query]. Companies often publish high-level AI principles or corporate ESG reports, but detailed metrics on the bias levels of specific models, the outcomes of internal ethical reviews, the energy consumed per AI query, or the full lifecycle emissions of AI hardware are rarely disclosed publicly. This opacity makes it difficult to gauge genuine progress, compare entities meaningfully, or hold organizations truly accountable for their impacts.

This report aims to address this challenge by providing a comparative analysis and ranking of 25 key entities based on their *demonstrable* commitment to ethical standards and environmental sustainability, focusing specifically on **transparency and accountability** as revealed through publicly available information. Due to the data limitations, this assessment relies on proxies: the clarity and substance of stated

principles, the transparency of governance structures, the openness and effectiveness of responses to incidents, the specificity of environmental reporting related to AI, and the evidence of concrete actions and commitments. The 25 entities evaluated are: AWS/Amazon, NVIDIA, AMD, Intel, Huawei, Cerebras Systems, Graphcore, Mistral AI, Cohere, AI2 (Allen Institute for AI), Mila (Quebec AI Institute), Vector Institute, Amii, BAAI (Beijing Academy of Artificial Intelligence), Alibaba Cloud, Google/Alphabet, Microsoft, Meta, IBM, Anthropic, OpenRouter, Groq, DeepSeek, OVHCloud, and OpenAI [Query]. The objective is to provide stakeholders with a structured, evidence-based evaluation of how these influential players are navigating the ethical and environmental dimensions of AI, while explicitly acknowledging the limitations imposed by the current state of public disclosure.

II. Assessment Framework: Measuring Ethical & Environmental Accountability

Rationale: Given the scarcity of standardized and verified performance data, a structured assessment framework is necessary to enable a systematic comparison of the selected entities. This framework focuses on evaluating the *transparency* and *accountability* demonstrated by each entity through its public disclosures and actions. It assesses the extent to which companies openly communicate their ethical principles, governance processes, bias mitigation efforts, environmental impacts specific to AI, and responses to failures. It is crucial to understand that this framework primarily measures the *willingness and ability to report and be accountable*, which serves as a necessary, albeit imperfect, proxy for underlying ethical and environmental performance [Query]. High scores indicate greater transparency and demonstrable accountability based on public evidence, while low scores reflect opacity or a lack of publicly available evidence regarding these aspects.

Scoring Rubric Dimensions: The assessment utilizes a rubric based on six key dimensions, applied consistently across all 25 entities. Scores are assigned on a qualitative scale (e.g., Low, Medium, High) based on the evidence gathered for each dimension:

- **A. Ethical Governance Transparency & Consistency:**
    - *Measures:* This dimension evaluates the clarity, public availability, and perceived stability of an entity's stated AI ethics principles. It assesses the transparency regarding internal governance structures established to oversee AI ethics, such as dedicated committees, review boards, or responsible AI offices, including details on their mandate, composition, and public reporting on their activities or influence.[2] It also considers evidence of consistency between these stated principles and governance structures and the entity's documented actions, such as product development decisions or responses to ethical challenges.
    - *High Score Indicators:* Clear, comprehensive, and easily accessible AI principles. Detailed public information on governance structures, their specific mandates, operational processes, and influence or outputs. Evidence demonstrating alignment between stated principles and company actions (e.g., documented impact of ethics reviews).
- **B. Ethical Accountability & Responsiveness:**
    - *Measures:* This dimension assesses the entity's track record in acknowledging

and addressing ethical failures, controversies, or identified biases in its AI systems.[3] It evaluates the transparency of reporting on such incidents and their resolution. Crucially, it looks for demonstrable actions taken in response to ethical concerns, whether raised internally or externally, such as specific policy changes, product modifications or withdrawals, legal settlements, or establishment of redress mechanisms.[3] The existence and accessibility of channels for stakeholder input, feedback, or seeking redress are also considered.[57]
    - *High Score Indicators:* Proactive, timely, and transparent acknowledgment of ethical issues or failures. Clear communication regarding remedial actions undertaken. Verifiable evidence of changes in practices, policies, or products resulting from ethical concerns. Robust, easily accessible mechanisms for stakeholder feedback and redress.
- **C. Bias & Fairness Transparency:**
    - *Measures:* This dimension focuses specifically on transparency related to algorithmic bias and fairness. It evaluates the entity's openness about the inherent risks of bias in its AI models and systems.[3] It assesses the provision and support of publicly accessible tools, frameworks, or techniques designed to help users or developers detect and mitigate bias.[34] Furthermore, it examines the level of transparency regarding the datasets used for training major AI models, including information about data composition, sourcing, and potential embedded biases.[68]
    - *High Score Indicators:* Explicit acknowledgment of bias risks in official documentation, model cards, or reports. Publicly available, well-documented, and supported tools or frameworks for fairness assessment and mitigation. Detailed, accessible information on training data sources, composition, and known limitations for key AI models.
- **D. Environmental Reporting Specificity & Verification:**
    - *Measures:* This dimension evaluates the granularity and credibility of the entity's environmental reporting, specifically concerning the impacts of its AI activities. It assesses the level of detail provided on AI's energy consumption, water usage, greenhouse gas (GHG) emissions (including Scope 1, 2, and particularly Scope 3 related to hardware manufacturing and supply chain), and hardware lifecycle impacts, as distinct from overall corporate environmental data.[1] It considers the use of standardized reporting frameworks (like the GHG Protocol) specifically applied to AI operations. Finally, it looks for evidence of independent, third-party verification or assurance for the reported environmental data *related to AI*.
    - *High Score Indicators:* Dedicated sections or metrics in sustainability reports

focusing on AI's environmental footprint. Data broken down by AI training versus inference, per AI task, or per hardware unit. Application of recognized standards (e.g., GHG Protocol Scopes) to AI operations. Explicit mention of AI-related environmental data within the scope of third-party assurance statements.

- **E. Environmental Action & Commitment:**
  - *Measures:* This dimension assesses tangible actions taken and specific commitments made to reduce AI's environmental impact, going beyond research or general corporate goals. It looks for evidence of deployed "Green AI" initiatives, such as documented use of renewable energy specifically procured for or allocated to AI workloads [17], deployment and promotion of energy-efficient AI-specific hardware (like GPUs, TPUs, accelerators) and software/models [6], and water efficiency measures specifically implemented for data centers hosting AI workloads.[5] It also evaluates the existence of clear, time-bound commitments to reduce the *specific* environmental footprint of the entity's AI operations or products.
  - *High Score Indicators:* Concrete, documented examples of deployed efficiency measures or renewable energy sourcing specifically linked to AI operations. Specific, measurable, achievable, relevant, and time-bound (SMART) targets for reducing AI's environmental impact (e.g., energy consumption per inference, water usage effectiveness for AI data centers, emissions from AI hardware lifecycle).
- **F. Overall Transparency Willingness:**
  - *Measures:* This provides a holistic, qualitative assessment combining the transparency demonstrated across all the preceding ethical and environmental dimensions. It considers how readily and clearly the entity volunteers information versus requiring external pressure, regulatory mandates, or investigative reporting to elicit disclosures. The accessibility, clarity, and comprehensiveness of the information provided are key factors.
  - *High Score Indicators:* A consistent pattern of proactive, clear, comprehensive, and easily accessible disclosures across multiple ethical and environmental domains related to AI. Evidence of engaging openly with critics or concerns.

Scoring Scale: A three-tier qualitative scale (Low, Medium, High) is used for each dimension.

- **High:** Demonstrates significant transparency and accountability through clear, detailed public disclosures, evidence of robust processes, and proactive engagement on the dimension's issues.
- **Medium:** Provides some relevant public information but lacks detail, consistency,

or verifiable evidence. Disclosures may be partial, difficult to find, or primarily high-level statements.
- **Low:** Minimal or no relevant public information available on the dimension. Disclosures are absent, vague, or purely aspirational without supporting evidence.

Weighting for Overall Rank: To generate an overall ranking that reflects the query's emphasis, the dimensions are weighted as follows, reflecting a focus on demonstrable accountability and the challenging areas of AI-specific environmental impact and governance:

- Ethical Governance Transparency & Consistency (A): 15%
- Ethical Accountability & Responsiveness (B): 20%
- Bias & Fairness Transparency (C): 15%
- Environmental Reporting Specificity & Verification (D): 25%
- Environmental Action & Commitment (E): 15%
- Overall Transparency Willingness (F): 10% Total: 100%

Data Collection Methodology: Information was systematically gathered for each of the 25 entities from publicly accessible sources, including official company websites (AI principles pages, ethics reports, sustainability/environmental reports, annual reports, governance sections, blogs, press releases), public statements, policy submissions, documented incident reports, reputable third-party sources (investigative journalism, academic research, NGO reports like those from AI Now Institute [138], Greenpeace [142], Amnesty Tech [145], and regulatory filings), and information regarding participation in multi-stakeholder initiatives (e.g., Frontier Model Forum [29], AI Alliance [149]).

III. Comparative Analysis: Performance Across Key Dimensions

This section analyzes the performance of the 25 target entities across the six dimensions defined in the assessment framework. It highlights leaders, laggards, and notable trends, supported by evidence from public sources.

A. Ethical Governance Transparency & Consistency Analysis

The establishment of clear ethical principles and transparent governance structures is fundamental for building public trust and guiding the responsible development and deployment of AI technologies. This dimension assesses the extent to which entities publicly articulate their AI principles, provide visibility into their internal governance mechanisms (such as ethics boards or review processes), and demonstrate consistency between their stated commitments and observable actions. The analysis reveals a wide spectrum of approaches, ranging from detailed public disclosures by major technology firms and some AI labs to minimal public information from smaller startups and certain non-Western companies. Research institutes typically publish ethical guidelines pertinent to research conduct.

- **Detailed Analysis & Examples:**
    - **Leaders in Governance Transparency:** Companies like **Microsoft**, **Google (Alphabet)**, **IBM**, and **Anthropic** stand out for providing relatively detailed public information about their AI principles and governance frameworks.
        - **Microsoft** outlines six core principles (Fairness, Reliability & Safety, Privacy & Security, Inclusiveness, Transparency, Accountability) [29] and details its governance structure, including the Office of Responsible AI (ORA), the Responsible AI Standard, the Aether Committee (providing research expertise), and the Responsible AI Council (co-led by senior executives).[29] They also publish a Responsible AI Transparency Report.[29]
        - **Google (Alphabet)** published its AI Principles in 2018 [92] and provides annual Responsible AI Progress Reports.[92] Their governance includes these principles, a Frontier Safety Framework for advanced models, content safety policies, application-specific frameworks, pre- and post-launch reviews involving leadership, and model documentation like Model Cards.[92] While corporate governance guidelines exist [45], specific details on the AI ethics review *process* influence might be less explicit than Microsoft's structure.
        - **IBM** emphasizes its Principles for Trust and Transparency (Augment Human Intelligence, Data Ownership, Transparency/Explainability) and Pillars of Trust (Explainability, Fairness, Robustness, Transparency,

Privacy).[32] It highlights the role of its AI Ethics Board, established in 2019, which provides governance, ensures consistency with values, reviews products, and advances trustworthy AI.[32]

- **Anthropic** promotes its "Constitutional AI" approach [150] and publishes a detailed Responsible Scaling Policy (RSP).[48] Their Voluntary Commitments page outlines internal governance frameworks, risk assessment processes involving internal and external reviews, and alignment with standards like ISO/IEC 42001.[48]
- **OpenAI** operates under a unique capped-profit structure governed by a non-profit board with a mission focused on safe and beneficial AGI.[54] Recent changes include the formation of a dedicated Safety and Security Committee (SSC) as an independent Board oversight committee with authority over model launches.[54] This structure, while complex and subject to public scrutiny, provides a degree of transparency about its governance intent.

- **Hardware Companies: NVIDIA** [47], **AMD** [37], and **Intel** [31] articulate AI principles focused on aspects like privacy, safety, transparency, and non-discrimination. They mention internal structures like Intel's multidisciplinary advisory councils [31] and AMD's Responsible AI Council.[37] However, public details on how these structures concretely influence product development or specific governance *outcomes* are generally less available compared to the AI/cloud leaders. **Cerebras Systems** emphasizes ethical development primarily through its participation in collaborative initiatives like the AI Alliance [149] and provides less detail on its internal governance structures.[66] **Graphcore** publishes a Supplier Ethics and Compliance Policy [155] but offers limited public information on its specific AI governance framework.[139]
- **Other AI Labs/Startups: Cohere** outlines its Secure AI Frontier Model Framework, detailing a five-component approach to safety and security focused on practical enterprise risks [64] and mentions a Responsibility Council [156], but provides less public detail on the council's mandate or broader governance processes compared to leaders. **Mistral AI** emphasizes an open-source approach but has limited public documentation on its formal internal AI governance principles or structures.[2] **DeepSeek** [161] and **Groq** [67] appear to have minimal public information regarding their AI governance frameworks. **OpenRouter**, functioning as an API gateway [72], adheres to the principles of the models it routes to but has less need for its own comprehensive model governance structure.[39]
- **Cloud Providers: AWS (Amazon)** outlines core dimensions of responsible AI (Fairness, Explainability, Privacy/Security, Safety, etc.) and offers tools and

resources [34], but provides less explicit detail on its internal *governance board* structure compared to Microsoft or IBM. Its governance appears more integrated into product development and security processes.[59] **Alibaba Cloud** mentions security and responsible AI practices in product descriptions [163] and its founder has called for responsible AI [164], but official, detailed AI principles or governance structures are not readily found in the provided English-language materials.[35] **OVHcloud** details its corporate board structure, including a strategic and ESG committee [61], but specific AI governance principles or processes are not highlighted.[36] **Huawei** discusses AI governance concepts, has published white papers [50], and mentions compliance officers and subsidiary boards [50], but a clearly articulated set of public AI principles comparable to Google or Microsoft is less evident.[30]

    - **Research Institutes: Mila** [49], **AI2 (Allen Institute for AI)** [68], **Vector Institute** [69], **Amii** [70], and **BAAI (Beijing Academy of Artificial Intelligence)** [71] generally publish principles or guidelines focused on responsible research, ethical conduct, collaboration, and societal benefit. Mila has specific Rules of Conduct for partners.[90] Vector Institute outlines clear AI Trust & Safety Principles.[170] Amii details its Principled AI Strategy with core principles (Fairness, Privacy, Accountability).[171] BAAI published the Beijing AI Principles covering aspects like "Do Good" and "Responsibility".[172] Their governance structures typically relate to institute oversight rather than the governance of deployed commercial products. AI2 emphasizes transparency through its open model initiatives like OLMo.[68]

- **Governance Structures: Visibility vs. Influence:** A recurring observation is that while numerous companies publicly state the existence of AI ethics boards, councils, or policies [29], the actual influence and decision-making power of these bodies often remain opaque. Public reports rarely detail specific instances where these governance structures have altered product roadmaps, vetoed deployments, or mandated significant changes based on ethical concerns. The high-profile governance changes and debates at OpenAI [54], while tumultuous, offered an unusual, albeit indirect, view into the potential impact and contested nature of AI governance mandates. Without more transparency regarding the *processes* and *outcomes* of these internal bodies, it is challenging to assess whether they function as substantive checks on development or primarily as symbolic gestures ("ethics washing").[52] The public mandate and reporting detail thus become crucial proxies for evaluating the potential substance of these governance claims.
- **Consistency Between Principles and Practice:** Evaluating the consistency between stated principles (like fairness or safety) and documented company

actions or product behaviors is essential but difficult with only public data. High-profile incidents involving bias [3], safety failures, or privacy concerns [79] often appear to contradict companies' espoused values.[29] For example, Google's stated AI Principles [92] faced scrutiny following the Gemini image generation bias incident [73], and Microsoft's fairness principle [29] was challenged by findings of bias in its facial recognition technology.[74] This apparent gap underscores the significant challenges in operationalizing high-level principles effectively across large organizations and complex AI systems. Consequently, a company's public response and demonstrable actions following such incidents become critical indicators of its commitment to accountability and the practical application of its principles.

B. Ethical Accountability & Responsiveness Analysis

This dimension moves beyond stated principles and governance structures to examine how entities actually respond when ethical issues arise. It focuses on the public acknowledgment of failures, transparency about incidents and their resolution, evidence of concrete remedial actions, and the availability of mechanisms for stakeholder input and redress. This provides a practical measure of whether ethical commitments translate into tangible accountability.

- **Detailed Analysis & Examples:**
  - **Incident Handling and Acknowledgment:** The response to public controversies provides a key test of accountability.
    - **Google** publicly acknowledged the issues with Gemini's image generation, paused the feature, and provided an explanation attributing it to over-correction in bias mitigation efforts.[73] This contrasts with some earlier industry responses to facial recognition bias that were perceived as more dismissive.[74]
    - **Microsoft**, when faced with Dr. Buolamwini's findings on facial recognition bias, initially saw issues persist even after updates, though IBM was noted as being more responsive in that instance.[74] Microsoft has since invested heavily in responsible AI tooling and frameworks.[93]
    - **Meta** responded to the controversy over its AI-generated accounts ('Liv') by deactivating them and framing the initiative as an internal test, potentially downplaying the systemic issues raised about bias and lack of diversity in creation.[78] Meta has also faced significant criticism and investigations regarding its platform's role in amplifying harmful content and impacting human rights, particularly concerning the Rohingya in Myanmar [146] and Tigrayans in Ethiopia.[146]

        - **OpenAI** has faced lawsuits alleging data misuse [80] and internal concerns regarding the prioritization of safety.[56] Its response involved legal defenses and the formation of the Safety and Security Committee [55], representing a structural response to accountability demands.
        - **Anthropic** has publicly discussed research on "alignment faking" [76], framing it within its ongoing safety research, demonstrating a degree of transparency about potential model risks.
        - **Huawei** faced allegations related to lobbying practices in the EU, distinct from AI product failures but relevant to corporate ethical conduct.[77]
        - **Cerebras Systems** faced scrutiny over its relationship with G42 due to G42's alleged ties with China, prompting Cerebras to emphasize compliance and US interests.[87]
        - For many smaller entities (e.g., Graphcore, Groq, DeepSeek, OpenRouter) and some research institutes, there is a lack of publicly documented major AI ethical controversies. This may be due to lower public scrutiny, a different operational focus (hardware vs. public-facing models), or genuinely fewer incidents, but the lack of reporting makes assessment difficult.
    - **Redress and Stakeholder Input:** Some companies explicitly mention mechanisms for reporting issues or providing feedback. Anthropic details its Responsible Disclosure Policy, bug bounty programs, and a user safety email address.[48] Microsoft provides various channels for feedback through its platforms and documentation.[57] AWS encourages responsible use and provides resources but specific public redress channels are less prominent.[34] Mila's Rules of Conduct for Partners include specific enforcement actions Mila can take, acting as a form of accountability within its ecosystem.[90] The general accessibility and effectiveness of these mechanisms for external stakeholders are often hard to gauge from public information alone.
    - **Proactive Accountability Measures:** Some initiatives suggest a more proactive stance. Anthropic's Responsible Scaling Policy [48] and Constitutional AI [150] aim to embed safety and ethical constraints from the outset. Microsoft's Copilot Copyright Commitment offers indemnification to customers, proactively addressing a potential legal and ethical risk.[29] OpenAI's establishment of the Safety and Security Committee, particularly its mandate to oversee model launches, represents a structural commitment to accountability.[55] IBM's long-standing AI Ethics Board serves a similar function.[32]
- **The Recurrence of Public Failures:** Observing the sequence of events around several high-profile AI controversies—product launch, public criticism regarding

bias or safety, company apology/explanation, and subsequent adjustments—suggests a recurring pattern.[73] This pattern implies that pre-deployment testing, internal ethical reviews, and risk assessments may sometimes be insufficient or perhaps deprioritized relative to the speed of innovation and market pressures. This leads to a cycle where public failures become a de facto part of the development and refinement process, placing the burden of identifying harms onto users and the public, rather than being fully addressed proactively within the developing organization. This highlights the critical importance of robust internal governance actually influencing development *before* release.

- **Framing Accountability:** The language and framing used by companies when responding to ethical failures are significant indicators of their approach to accountability. Responses often emphasize the complexity of AI, unintended consequences, or the experimental nature of the technology [[73] (Google framing), [76] (Anthropic framing), [85] (Cohere framing risk), [73] (Google framing), [78] (Meta framing)]. While acknowledging issues is a necessary step, framing them purely as technical glitches or unavoidable side effects of innovation can sometimes serve to deflect from deeper accountability for potential systemic flaws in data sourcing, model design, testing protocols, or the adequacy of internal governance oversight. Public relations and legal considerations invariably shape these public narratives, making it essential to critically evaluate the substance of the response beyond the initial acknowledgment.

C. Bias & Fairness Transparency Analysis

Algorithmic bias remains a pervasive and critical ethical challenge in AI development and deployment, potentially leading to discriminatory outcomes in sensitive areas like hiring, lending, healthcare, and law enforcement.[3] This dimension assesses transparency specifically related to bias and fairness, examining the openness about risks, the provision of tools for assessment and mitigation, and transparency concerning training data.

- **Detailed Analysis & Examples:**
  - **Risk Acknowledgment:** Acknowledging the risk of bias is a prerequisite for addressing it. Many of the assessed entities, particularly the larger tech companies and AI labs, include fairness or non-discrimination as core AI principles and publicly acknowledge the potential for bias in their systems.[29] The depth varies from general statements to more specific discussions, sometimes prompted by public incidents like Google's Gemini controversy [73] or the historical issues with facial recognition bias across multiple vendors.[74]

Huawei explicitly mentions checking outputs for bias and ensuring diverse development teams.[50]

- **Tools and Frameworks:** Several leading companies provide tools aimed at helping developers assess and mitigate bias, often releasing them as open-source projects:
  - **Microsoft** offers the Fairlearn toolkit (Python) for assessing and mitigating group fairness harms, along with visualization tools in its Responsible AI Dashboard.[93]
  - **IBM** developed AI Fairness 360, an open-source library with a comprehensive set of metrics and mitigation algorithms.[94]
  - **Google** created the What-If Tool for probing model behavior and mentions using evaluation metrics for representational harms in models like Gemma.[94]
  - **Meta** is associated with the Fairness Flow tool (mentioned in external sources [94], though detailed public documentation seems limited in the provided materials) and has an Inclusive AI program.[94] The Aequitas framework, while not directly Meta's, focuses on similar goals.[173]
  - **AWS** provides tools like Amazon SageMaker Clarify for bias detection and explainability.[34] The availability, documentation quality, ease of use, and ongoing support for these tools vary, impacting their practical utility for developers and researchers.
- **Dataset Transparency:** This remains a significant area of opacity, especially for large, proprietary foundation models. While some Model Cards (e.g., Google's for Gemma [102], Anthropic's for Claude 3 [48]) provide high-level descriptions of training data or list datasets used for evaluation, comprehensive details about the composition, sourcing methods, cleaning processes, and known biases within the massive pre-training corpora are generally not disclosed. This lack of transparency makes it difficult for external parties to independently assess potential biases originating from the data itself. **AI2** stands out by emphasizing openness with its OLMo models and the OLMoTrace tool, specifically designed to allow tracing model outputs back to the training data.[68] Companies like **DeepSeek** provide very little public information about their training data.[161] **Huawei** mentions the importance of representative datasets [50] but doesn't detail its own practices for major models.

- **Gap Between Tool Availability and Implementation:** While the development and release of fairness toolkits by major players like Microsoft, IBM, and Google represent positive steps towards enabling bias mitigation [93], the mere existence of these tools does not guarantee their effective implementation or the prevention

of biased outcomes in practice. Public reporting rarely includes metrics on the internal adoption rates of these tools, the findings from internal fairness audits, or the measured impact of mitigation techniques applied during development. The recurrence of bias-related incidents at companies that also provide fairness tools [73] suggests a potential gap between the availability of resources and their consistent, effective application throughout the development lifecycle. Greater transparency regarding the *use* and *impact* of these tools internally is needed to assess their real-world effectiveness in mitigating bias.

- **The Open Source Transparency Dynamic:** The rise of powerful open-source models like Meta's Llama [83], Mistral's models [38], and AI2's OLMo [68] introduces a complex dynamic regarding transparency and risk. Open models offer potentially greater visibility into model architecture and, in some cases (like OLMo), training data, facilitating external scrutiny, research, and identification of biases.[68] This aligns with principles of scientific openness. However, this very openness can also make biases more readily apparent, leading to public controversies.[83] Furthermore, the open availability allows for potential misuse or deployment by actors who may lack the resources or inclination to implement necessary safety guardrails.[39] Conversely, developers of closed, proprietary models (like OpenAI's GPT-4 or Anthropic's Claude) can maintain tighter control over deployment and safety measures but offer significantly less transparency into the model's inner workings and training data, making independent verification of their claims difficult. This suggests that neither open nor closed approaches inherently guarantee better ethical outcomes; the robustness of the governance, testing, and accountability practices surrounding the model, regardless of its accessibility, remains paramount.

D. Environmental Reporting Specificity & Verification Analysis

Evaluating the environmental sustainability of AI necessitates specific data on its resource consumption (energy, water) and emissions, distinct from general corporate environmental reporting. This dimension assesses the level of detail and credibility provided by entities regarding AI's specific environmental footprint and whether this information is independently verified. The overall finding is that AI-specific environmental reporting is nascent, often lacking granularity and independent verification, making accurate assessment across the industry extremely challenging.

- **Detailed Analysis & Examples:**
  - **Aggregated Corporate Data vs. AI-Specific Data:** The vast majority of entities with public sustainability reporting, particularly large corporations like **Google** [4], **Microsoft** [24], **Meta** [7], **IBM** [108], **Amazon/AWS** [9], **NVIDIA** [113], **AMD** [124],

and **Intel** [106], publish annual sustainability or corporate responsibility reports. These typically include overall corporate GHG emissions (often broken down by Scope 1, 2, and 3), total energy consumption, water withdrawal/consumption, and waste generation. However, these reports rarely disaggregate the specific contribution of AI research, development, training, and inference activities from the company's total environmental footprint. Attributing a precise portion of data center energy use or Scope 3 hardware emissions solely to AI is generally not possible based on current public disclosures.

- **AI Mentions within Environmental Reports:** While specific AI footprint data is scarce, corporate sustainability reports increasingly *acknowledge* the growing impact of AI and data center expansion on overall environmental metrics.
  - **Microsoft's** 2024 report explicitly attributes the significant increase in its Scope 3 emissions (up 30.9% from 2020 baseline) primarily to the construction of more data centers for AI and the embodied carbon in hardware components.[24] It also notes the challenges AI infrastructure poses to meeting sustainability commitments.[24]
  - **Google's** 2024 report discusses building energy-efficient infrastructure for the AI era (e.g., Trillium TPUs) and links increased electricity consumption to data center growth.[23] It also notes a 13% year-over-year increase in total GHG emissions, driven partly by data center energy use.[16] Increased water consumption is also linked to data center cooling needs driven by AI expansion.[4]
  - **Amazon's** 2023 report mentions exploring AI applications for optimizing energy use and packaging.[118]
  - **NVIDIA's** reports emphasize the energy efficiency benefits of its GPUs for AI compared to CPUs and highlight their role in powering energy-efficient supercomputers.[114]
  - Reports from **Meta** [107] and **IBM** [108] discuss overall corporate goals but provide less specific linkage between AI and reported environmental metrics in the snippets reviewed.
- **Specific AI Footprint Data:** Publicly released, company-specific data on the environmental footprint *of AI itself* is extremely rare.
  - **Google** published a life-cycle assessment (LCA) for its TPU hardware, including estimates of manufacturing emissions and introducing a "compute carbon intensity" (CCI) metric, claiming a 3x improvement from TPU v4 to Trillium.[112] This represents a notable, albeit hardware-specific, instance of transparency.

        - Academic studies and third-party analyses have estimated the footprint of training specific models like OpenAI's GPT-3 [8] or Meta's Llama [22], or the energy cost per query for chatbots versus standard search.[4] However, these are typically external estimates, not official, regularly updated disclosures from the companies (except Google's TPU LCA).
        - Leading AI labs like **OpenAI**, **Anthropic**, **Cohere**, **Mistral AI**, and **DeepSeek** currently lack publicly available environmental or sustainability reports detailing their operational footprints (energy/water consumption, emissions from training/inference). Their impact is often indirectly captured within the reports of their cloud providers (AWS, Google Cloud, Azure) but is not separately disclosed.
        - Hardware companies (**NVIDIA**, **AMD**, **Intel**, **Cerebras**, **Groq**) report extensively on the *energy efficiency* of their products [6], but comprehensive, public LCAs covering manufacturing, use, and disposal for their AI-specific hardware (beyond Google's TPU study) are generally not available.
    - **Verification:** While many large companies obtain third-party assurance for their corporate sustainability reports [104], this assurance typically applies to the overall reported data (e.g., total Scope 1, 2, 3 emissions). It rarely, if ever, extends to verifying specific claims about AI's contribution to that footprint or the impact of AI-specific mitigation measures, unless explicitly stated within the assurance scope. Google's TPU LCA, for example, appears to be an internal study.[112]
- **The AI Environmental Data Deficit:** A clear pattern emerges: despite widespread acknowledgment of AI's significant and growing environmental impact—particularly concerning energy and water consumption driven by data centers [4]—there is a profound lack of specific, granular, and verified public reporting on this impact by the very companies developing and deploying these technologies. Most corporate sustainability reports [23] aggregate AI's footprint within broader corporate metrics, effectively obscuring the specific environmental cost of AI innovation and deployment. This transparency gap hinders the ability of stakeholders to accurately assess the sustainability of different AI approaches, compare companies' performance, and hold the industry accountable for its environmental responsibilities.
- **Efficiency Claims vs. Absolute Growth:** A common narrative, particularly from hardware manufacturers (NVIDIA, AMD, Intel, Cerebras, Groq) and cloud providers, focuses on improvements in *energy efficiency*—delivering more computational performance per unit of energy.[6] While these technological advancements are real and important, they are occurring alongside an

exponential increase in the *scale* of AI deployment. This rapid growth in demand for AI computation often outpaces efficiency gains, leading to a significant rise in *absolute* energy consumption, water usage, and associated emissions for the sector as a whole and for major players.[4] This phenomenon, akin to the Jevons paradox [15], means that focusing solely on efficiency metrics provides an incomplete and potentially misleading picture of AI's total environmental burden. True transparency requires reporting both relative efficiency improvements *and* the absolute environmental footprint attributable to AI activities.

E. Environmental Action & Commitment Analysis

This dimension evaluates the concrete actions entities are taking to mitigate the environmental impact of their AI operations and the specificity of their public commitments for future reductions. It looks beyond reporting to assess deployed initiatives and targeted goals related to AI sustainability.

- **Detailed Analysis & Examples:**
  - **Deployed Green AI Initiatives:** Several companies highlight specific actions aimed at improving the sustainability of AI infrastructure and operations:
    - **Google** emphasizes its development and use of energy-efficient Tensor Processing Units (TPUs) [23], reports high power usage effectiveness (PUE) in its data centers [23], uses AI itself to optimize data center cooling [25], and promotes AI for broader environmental solutions like flood prediction and efficient routing.[23]
    - **Microsoft** is developing new data centers designed for AI workloads that will consume zero water for cooling.[24] They use AI for water leak detection (FIDO) [24], require key suppliers to use 100% carbon-free electricity (CFE) for Microsoft goods/services by 2030 [24], and operate Circular Centers for hardware reuse and recycling.[134]
    - **Amazon (AWS)** uses AI to optimize packaging and monitor energy use in buildings [118] and promotes the efficiency of running workloads on AWS versus on-premises.[117]
    - **Huawei** developed the iCooling@AI solution for smart data center cooling to reduce PUE.[105]
    - **Hardware Companies (NVIDIA, AMD, Intel, Cerebras, Groq)** focus heavily on designing more energy-efficient chips (GPUs, AI accelerators) as a primary environmental action.[6] **Cerebras** highlights its partnership with Green AI Cloud, which utilizes renewable energy and waste heat reuse.[17] **Groq** positions its LPU technology as enabling efficient AI.[18]
    - **Research Institutes: Mila's** AI computing cluster incorporates

energy-efficient design with heat reuse for campus buildings.[174] **AI2**, **Vector Institute**, and **Amii** focus on AI *for* environmental applications like conservation, climate modeling, and resource optimization.[130]

- **AI-Specific Commitments:** Specific, measurable, time-bound commitments focused solely on reducing the environmental footprint of *AI operations* are rare. Most publicly stated goals relate to overall corporate operations:
  - Many large companies have goals to achieve 100% renewable energy matching for their operations by specific dates (e.g., Amazon/AWS reached this in 2023 [118], Microsoft aims for 100% CFE coverage by 2030 [24], Google aims for 24/7 CFE by 2030 [23], NVIDIA aims for 100% renewable electricity for offices/data centers by end of FY25 [128], AMD suppliers aim for 80% renewable sourcing by 2025 [124]).
  - Net-zero commitments for overall value chains are also common (e.g., Amazon by 2040 [118], Google by 2030 [23], Microsoft aiming for carbon negative by 2030 [24]).
  - While these corporate goals indirectly benefit AI operations hosted within them, they lack the specificity needed to track progress on mitigating AI's unique and rapidly growing footprint.
  - **AMD** has a specific goal to achieve a 30x energy efficiency improvement for processors and accelerators used in HPC and AI training between 2020 and 2025.[37] This is a notable AI-relevant efficiency target.
  - **Microsoft's** requirement for key suppliers to use 100% CFE by 2030 [24] is a significant commitment impacting the Scope 3 emissions associated with hardware manufacturing, which is highly relevant to AI's footprint.
- **Cloud Provider Tools for Customers:** Major cloud providers like **AWS** [117], **Google Cloud** [116], and **Azure** [58] offer tools (like the AWS Customer Carbon Footprint Tool) that allow their *customers* to estimate the carbon footprint associated with their cloud usage. While this empowers users, it is an indirect action by the provider regarding their own AI development footprint.

- **Addressing the Primary Impact Drivers:** Current environmental actions often focus on improving operational energy efficiency (Scope 2, PUE) and promoting hardware efficiency (product use phase). While valuable, these efforts may not sufficiently address the largest and most rapidly growing components of AI's environmental impact. As highlighted by Microsoft's report [24], Scope 3 emissions, particularly the embodied carbon in constructing data centers and manufacturing vast quantities of specialized hardware (chips, servers, racks), are becoming dominant drivers of the overall footprint. Furthermore, the sheer *scale* of AI deployment is driving absolute increases in energy and water demand, potentially negating relative efficiency gains.[10] Truly comprehensive environmental action

needs to tackle these Scope 3 emissions through supply chain decarbonization (like Microsoft's supplier CFE requirement [24]), circular economy principles for hardware [24], and potentially strategies to manage demand growth, alongside continued efficiency improvements.

- **AI *for* Sustainability Narrative:** A common theme across many entities is the promotion of AI as a tool *for* achieving environmental sustainability goals.[1] Examples include using AI for climate modeling, optimizing energy grids, discovering sustainable materials, improving agricultural efficiency, and monitoring deforestation or illegal fishing. While these applications hold genuine promise, this narrative can sometimes overshadow or implicitly justify the significant environmental cost incurred in developing and deploying the AI tools themselves. A balanced and transparent assessment requires acknowledging both the potential environmental benefits enabled *by* AI applications and the environmental footprint generated *from* creating and running those AI systems. Without clear data on the latter [Insight 7], it is difficult to objectively weigh the net environmental impact.

F. Overall Transparency Willingness Synthesis

This final dimension provides a qualitative synthesis of each entity's overall posture towards transparency and openness across the ethical and environmental domains assessed. It considers the proactiveness, clarity, accessibility, and comprehensiveness of disclosures, moving beyond individual metrics to evaluate the general willingness to engage openly with stakeholders about AI's impacts and governance.

- Analysis:
  A clear pattern emerges where transparency levels often correlate with company size, type, market position, and geographic location.
  - **High Transparency Willingness:** Entities like **Google/Alphabet**, **Microsoft**, **IBM**, and **Anthropic** generally demonstrate higher transparency willingness. They publish dedicated reports on responsible AI [29], detail their principles and governance structures [29], offer fairness tools [93], acknowledge incidents (to varying degrees) [42], and publish extensive corporate sustainability reports that, while often lacking AI specifics, provide a baseline of environmental disclosure.[24] **OpenAI**, despite past criticisms, has increased transparency around its governance structure and safety committee.[54] **AI2** shows strong willingness through its open model and traceability initiatives.[68] Research Institutes (**Mila**, **Vector**, **Amii**, **BAAI**) generally exhibit transparency aligned with academic norms regarding principles and research goals.[49]

- **Medium Transparency Willingness:** This category includes major hardware vendors like **NVIDIA**, **AMD**, and **Intel**. They publish principles [31] and corporate responsibility reports detailing product efficiency and operational sustainability.[31] However, transparency regarding internal AI governance *influence* and AI-specific environmental footprints (beyond product efficiency) is less pronounced than the leaders. **AWS (Amazon)** provides responsible AI guidance and tools [34] and corporate sustainability reports [118], but has faced criticism for transparency regarding data center energy sources and lacks the detailed governance structure disclosures of Microsoft or IBM.[142] **Meta** publishes sustainability reports [107] and has fairness initiatives [94], but faces scrutiny over platform impacts and transparency regarding algorithmic harms.[78] **Cohere** [64] and **Mistral AI** [2] provide some information but generally less comprehensive public documentation on governance and environmental impact compared to larger players.
- **Low Transparency Willingness:** This group tends to include smaller or more specialized hardware companies like **Cerebras Systems** [17] and **Graphcore** [10], newer AI startups like **DeepSeek** [15] and **Groq** [18], and API aggregators like **OpenRouter** [20], for whom detailed public reporting on ethics governance and environmental footprint is minimal or non-existent based on the reviewed sources. Non-Western companies like **Huawei** [30] and **Alibaba Cloud** [35] publish corporate sustainability reports [104] but provide less accessible or detailed information specifically on AI ethics governance frameworks compared to Western leaders, potentially reflecting different regulatory environments and corporate communication norms. **OVHcloud** provides corporate governance details [61] and discusses sustainability [123] but lacks specific AI-focused transparency initiatives in the reviewed material.

- **Transparency as a Strategic Decision:** The observed variations strongly suggest that the level of public transparency is not solely determined by an entity's inherent ethical or environmental performance, but is significantly influenced by strategic considerations. Large, publicly traded Western companies (Google, Microsoft, Meta, IBM, NVIDIA, Intel, AMD) face greater pressure from investors, regulators (e.g., EU AI Act [29]), NGOs [138], media, and employees to disclose information on ESG and responsible AI practices. They also possess greater resources to dedicate to reporting and governance functions. Smaller startups may lack these resources or prioritize rapid product development over comprehensive public reporting. Hardware-focused companies may concentrate their transparency efforts on product specifications (like energy efficiency) rather than broader ethical governance. Non-Western companies operate under different regulatory landscapes and stakeholder expectations (e.g., China [167]),

which shapes their disclosure priorities. Research institutes prioritize academic openness and collaboration but may have limited capacity or mandate for detailed operational environmental reporting comparable to large corporations. This context dependence means that transparency levels reflect a complex interplay of external pressures, internal priorities, company maturity, business model, and geographic location, rather than being a pure measure of virtue.

IV. Rankings & Key Findings

This section presents the rankings derived from the assessment framework, both for individual dimensions and overall. It provides concise justifications for each entity's ranking based on the evidence analyzed in Section III and discusses overarching patterns observed across the AI landscape regarding transparency and accountability.

Presentation of Rankings:

The following tables summarize the rankings based on the analysis of publicly available information as of [Current Date]. Entities are ranked within each dimension and overall based on the scoring (Low/Medium/High translated to a numerical equivalent for ranking, e.g., Low=1, Medium=2, High=3) and the weighting scheme defined in Section II. *Note: Ties in scores result in shared ranks.*

Table 1: Ranking by Ethical Governance Transparency & Consistency

Rank	Entity	Score	Justification Summary
1	Microsoft	High	Detailed public principles, governance structure (ORA, Aether, Council), RAI Standard, Transparency Reports. [29]
2	IBM	High	Clear principles/pillars, public info on AI Ethics Board mandate & role, extensive documentation. [32]
3	Anthropic	High	Constitutional AI, detailed Responsible Scaling Policy,

			Voluntary Commitments outlining governance. [48]
4	Google/Alphabet	High	Public AI Principles, annual reports, Frontier Safety Framework, leadership reviews detailed. [92]
5	OpenAI	Medium	Unique non-profit governance, Safety Committee details public, but structure evolved post-controversy. [54]
6	Vector Institute	Medium	Clear AI Trust & Safety Principles published, focus on research governance. [170]
6	Amii	Medium	Public Principled AI Strategy with core principles, focus on research/industry application. [171]
7	Mila	Medium	Rules of Conduct for partners, focus on research ethics and collaboration. [49]
8	AI2 (Allen Institute)	Medium	Emphasis on open models (OLMo) & traceability (OLMoTrace) for transparency, research focus. [68]

9	BAAI	Medium	Published Beijing AI Principles covering key ethical themes. [172]
10	NVIDIA	Medium	Public principles, mentions internal structures, less detail on process influence. [47]
11	AMD	Medium	Public principles, Responsible AI Use Policy, mentions RAI Council, less detail on process. [37]
12	Intel	Medium	Public principles, mentions advisory councils, less detail on process influence. [31]
13	Cohere	Medium	Secure AI Framework outlined, mentions Responsibility Council, less detail on overall governance. [64]
14	AWS/Amazon	Medium	Responsible AI dimensions stated, tools offered, less explicit detail on central governance board. [34]
15	Meta	Medium	Principles exist (via Human Rights Policy link), Responsible AI teams mentioned, governance less centralized/clear. [52]

16	Huawei	Low	Discusses governance ideas, white papers, compliance officers, lacks clear public AI principles set. [30]
17	Cerebras Systems	Low	Mentions ethical development in alliances, minimal public info on internal AI governance. [66]
18	Mistral AI	Low	Limited public documentation on formal AI principles or internal governance structures. [2]
19	OVHCloud	Low	Corporate governance detailed, but AI-specific principles/processes not highlighted. [36]
20	Alibaba Cloud	Low	Responsible AI mentioned in product context, founder statements, lacks official principles doc. [35]
21	Graphcore	Low	Supplier ethics policy exists, minimal public info on AI-specific governance. [139]
22	Groq	Low	Minimal public information on AI governance or ethical principles. [67]

23	DeepSeek	Low	Minimal public information on AI governance or ethical principles. [161]
24	OpenRouter	Low	Functions as aggregator, minimal need/disclosure for own model governance. [39]

Table 2: Ranking by Ethical Accountability & Responsiveness

Rank	Entity	Score	Justification Summary
1	Anthropic	High	Proactive scaling policy, Constitutional AI, public discussion of risks (alignment faking), clear reporting channels. [48]
2	Microsoft	High	Acknowledged past issues (facial recognition), developed tools, copyright commitment, active policy engagement. [29]
3	Google/Alphabet	High	Acknowledged Gemini bias, paused feature, provided explanation, ongoing reporting. [73]
4	OpenAI	Medium	Formed Safety Committee in response to concerns/events,

			faces lawsuits, structural changes show responsiveness. [55]
5	IBM	Medium	Long-standing Ethics Board for review, sunset facial recognition tech citing concerns. [32]
6	Mila	Medium	Clear enforcement mechanisms outlined for partners engaging in misconduct. [90]
7	Meta	Low	Deactivated biased AI accounts but framed as experiment, faces ongoing criticism/investigations on platform harms. [78]
8	Cerebras Systems	Low	Responded to G42 concerns emphasizing compliance, but limited public track record on AI ethical incidents. [87]
9	NVIDIA	Low	Limited public record of acknowledging/responding to specific AI ethical controversies (e.g., bias concerns raised [75]).
10	AMD	Low	Responsible Use Policy prohibits harmful uses, but

			limited public record of responding to specific incidents. [43]
11	Intel	Low	Global Human Rights Principles allow action, but limited public record of responding to specific AI ethical incidents. [31]
12	Huawei	Low	Faced lobbying allegations, limited public info on response to AI-specific ethical issues. [77]
13	Cohere	Low	Discusses practical risks, limited public record of responding to specific incidents. [85]
14	Mistral AI	Low	Faces data privacy complaint, limited public response/action documented. [86]
15	AWS/Amazon	Low	Documented bias in past hiring tool [82], faces criticism on data center transparency/labor issues [142], limited AI incident response detail.
16	Vector Institute	Low	Focus on principles/research, limited public info on

			handling specific ethical incidents. [88]
17	Amii	Low	Focus on principles/framework, limited public info on handling specific ethical incidents. [88]
18	AI2	Low	Focus on open research, limited public info on handling specific ethical incidents. [88]
19	BAAI	Low	Focus on principles, limited public info on handling specific ethical incidents. [88]
20	Alibaba Cloud	Low	Limited public record of acknowledging/responding to specific AI ethical controversies (e.g., bias [83]).
21	Graphcore	Low	Minimal public information on AI ethical incident handling. [155]
22	Groq	Low	Minimal public information on AI ethical incident handling. [162]
23	DeepSeek	Low	Minimal public information on AI ethical incident handling. [161]

24	OpenRouter	Low	Minimal public information on AI ethical incident handling. [39]
25	OVHCloud	Low	Minimal public information on AI ethical incident handling. [36]

Table 3: Ranking by Bias & Fairness Transparency

Rank	Entity	Score	Justification Summary
1	Microsoft	High	Acknowledges bias risk, provides Fairlearn toolkit (open source), RAI Dashboard tools. [29]
2	Google/Alphabet	High	Acknowledges bias risk (esp. post-Gemini), provides What-If Tool, details evaluation metrics in reports. [73]
3	IBM	High	Acknowledges bias risk, provides AI Fairness 360 toolkit (open source). [32]
4	AI2	Medium	Strong emphasis on data transparency via OLMo/OLMoTrace, acknowledges bias issues in AI generally. [68]

5	Meta	Medium	Acknowledges bias risk, has Inclusive AI program, associated with Fairness Flow tool (external mentions). [94]
6	AWS/Amazon	Medium	Acknowledges bias risk, offers SageMaker Clarify, past issues with biased hiring tool documented. [34]
7	Anthropic	Medium	Acknowledges bias (societal impacts in commitments), focuses on Constitutional AI for alignment, less on specific tools/data. [48]
8	Huawei	Medium	Acknowledges bias risk, mentions need for representative datasets and diverse teams. [50]
9	Amii	Medium	Principled AI framework explicitly addresses Fairness/Non-Discrimination, references bias types. [171]
10	Vector Institute	Low	Principles mention democratic values/fairness, less focus on specific tools or dataset transparency. [170]

11	Mila	Low	Principles mention fairness, focus on research ethics, less on specific tools/data transparency. [90]
12	BAAI	Low	Principles mention fairness/diversity, less on specific tools/data transparency. [172]
13	NVIDIA	Low	Principles mention non-discrimination, less focus on public tools or dataset transparency. [152]
14	AMD	Low	Principles mention mitigating bias/inclusion, less focus on public tools or dataset transparency. [37]
15	Intel	Low	Principles mention equity/inclusion, less focus on public tools or dataset transparency. [31]
16	Cohere	Low	Acknowledges bias risk in general AI discussion, less focus on specific tools or dataset transparency. [85]
17	Mistral AI	Low	Limited public information on bias transparency efforts. [38]

18	Alibaba Cloud	Low	Limited public information on bias transparency efforts (incidents noted [83]).
19	Cerebras Systems	Low	Limited public information on bias transparency efforts. [149]
20	Graphcore	Low	Limited public information on bias transparency efforts. [155]
21	Groq	Low	Limited public information on bias transparency efforts. [162]
22	DeepSeek	Low	Limited public information on bias transparency efforts, notes lack of info on training data/guardrails. [161]
23	OpenRouter	Low	Limited public information on bias transparency efforts. [39]
24	OVHCloud	Low	Limited public information on bias transparency efforts. [36]
25	OpenAI	Low	Limited public information on specific bias tools or training dataset transparency beyond

			general statements. [54]

Table 4: Ranking by Environmental Reporting Specificity & Verification (AI Focus)

Rank	Entity	Score	Justification Summary
1	Google/Alphabet	High	Published specific TPU LCA with manufacturing emissions & CCI metric; Report links overall increases to AI/data centers. [23]
2	Microsoft	High	Report explicitly links Scope 3 increase to AI data center construction/hardware; Mentions AI impact challenges. [24]
3	NVIDIA	Low	Reports focus heavily on product efficiency, less on AI operational footprint or verified AI-specific data in CR reports. [113]
4	AMD	Low	Reports focus on product efficiency goal, less on AI operational footprint or verified AI-specific data. [124]
5	Intel	Low	Reports focus on product efficiency/corporate

			goals, less on AI operational footprint or verified AI-specific data. [6]
6	AWS/Amazon	Low	Corporate report mentions AI for efficiency, lacks specific AI footprint data; faces transparency criticism. [9]
7	Meta	Low	Corporate report outlines goals, lacks specific AI footprint data. [7]
8	IBM	Low	Corporate report outlines goals, uses AI for sustainability, lacks specific AI footprint data. [108]
9	Huawei	Low	Corporate report exists, mentions green innovation, lacks specific AI footprint data. [104]
10	Alibaba Cloud	Low	Mentions Energy Expert tool, lacks specific AI footprint data in public reports. [120]
11	Cerebras Systems	Low	Promotes hardware efficiency, partnership with Green AI Cloud, lacks own operational AI footprint report. [17]

12	OVHCloud	Low	Discusses data center efficiency (PUE, water cooling), lacks specific AI footprint report. [123]
13	OpenAI	Low	No public environmental report; footprint estimated externally or via cloud providers. [8]
14	Anthropic	Low	No public environmental report; footprint estimated externally or via cloud providers. [9]
15	Cohere	Low	No public environmental report; promotes AI for energy sector efficiency. [127]
16	Mistral AI	Low	No public environmental report; partnership with Veolia mentions resource efficiency. [1]
17	Graphcore	Low	No public environmental report; general industry impacts discussed externally. [10]
18	Groq	Low	No public environmental report; promotes LPU efficiency, Saudi partnership mentions sustainability. [18]

19	DeepSeek	Low	No public environmental report; external reports discuss potential efficiency. [15]
20	OpenRouter	Low	No public environmental report. [20]
21	Mila	Low	Mentions energy-efficient cluster design, no comprehensive environmental report. [174]
22	Vector Institute	Low	Holds events on Green AI, no comprehensive environmental report. [131]
23	Amii	Low	Promotes AI for industrial efficiency/sustainability, no comprehensive environmental report. [132]
24	AI2	Low	Focuses on AI for environmental applications, no comprehensive operational environmental report. [130]
25	BAAI	Low	Principles mention environment, no comprehensive operational environmental report.

			[19]

Table 5: Ranking by Environmental Action & Commitment (AI Focus)

Rank	Entity	Score	Justification Summary
1	Microsoft	High	Zero-water cooling for AI data centers, supplier CFE requirement, Circular Centers, AI for efficiency. [24]
2	Google/Alphabet	High	Efficient TPUs deployed, AI for data center cooling optimization, AI for sustainability apps, 24/7 CFE goal. [23]
3	AMD	Medium	Specific 30x energy efficiency goal for AI/HPC processors, supplier renewable energy goal. [37]
4	Cerebras Systems	Medium	Partnership with Green AI Cloud (renewables, heat reuse), promotes hardware efficiency. [17]
5	NVIDIA	Medium	Strong focus on GPU efficiency for AI, powers Green500 systems, corporate renewable goal. [114]
6	Intel	Medium	Focus on hardware

			efficiency, corporate renewable/net-zero goals. [6]
7	IBM	Medium	AI for efficiency/sustainability apps, corporate renewable/emissions goals. [108]
8	AWS/Amazon	Medium	Uses AI for operational efficiency, promotes cloud efficiency vs on-prem, corporate renewable goal met. [117]
9	Mila	Medium	Deployed energy-efficient cluster with heat reuse. [174]
10	Huawei	Low	Deployed iCooling@AI, corporate sustainability efforts, less specific AI commitments. [104]
11	Meta	Low	Corporate renewable/water/net-zero goals, uses AI for sustainability (concrete), less specific AI operational actions. [107]
12	OVHCloud	Low	Focus on data center water cooling/PUE, less specific AI actions/commitments

			. [123]
13	Groq	Low	Promotes LPU efficiency, Saudi partnership mentions sustainability, lacks concrete action details. [18]
14	DeepSeek	Low	External reports mention potential efficiency, lacks concrete action details. [15]
15	Cohere	Low	Promotes AI for energy sector efficiency, lacks concrete action details for own ops. [127]
16	Mistral AI	Low	Partnership mentions resource efficiency, lacks concrete action details for own ops. [1]
17	OpenAI	Low	Relies on cloud provider infrastructure, lacks public details on specific actions/commitments for own footprint. [8]
18	Anthropic	Low	Relies on cloud provider infrastructure, lacks public details on specific actions/commitments for own footprint. [9]

19	Graphcore	Low	Minimal public information on specific AI environmental actions/commitments. [10]
20	OpenRouter	Low	Minimal public information on specific AI environmental actions/commitments. [20]
21	Vector Institute	Low	Focus on research/convening (Green AI event), lacks details on operational actions. [131]
22	Amii	Low	Focus on research/application (AI for industrial efficiency), lacks details on operational actions. [132]
23	AI2	Low	Focus on research/application (AI for environment), lacks details on operational actions. [130]
24	BAAI	Low	Principles mention environment, lacks details on operational actions. [19]
25	Alibaba Cloud	Low	Offers Energy Expert tool, lacks details on specific AI

			operational actions/commitments.[120]

Table 6: Ranking by Overall Transparency Willingness

Rank	Entity	Score	Justification Summary
1	Microsoft	High	High transparency across governance, principles, tools, incident response (improving), and environmental reporting (acknowledging AI impact).
2	Google/Alphabet	High	High transparency on principles, reporting, governance details, TPU LCA, incident acknowledgment.
3	IBM	High	High transparency on principles, governance (Ethics Board), tools, corporate reporting.
4	Anthropic	High	High transparency via RSP, Constitutional AI, Voluntary Commitments detailing governance/risk assessment, proactive discussion of risks.
5	OpenAI	Medium	Increased transparency on governance/safety

			committee post-events, unique structure public, but lacks environmental/bias tool transparency.
6	AI2	Medium	High transparency through open models/data initiatives (OLMoTrace), research focus.
7	NVIDIA	Medium	Medium transparency via principles/CR reports/efficiency focus, lower on governance influence/AI-specific footprint.
8	AMD	Medium	Medium transparency via principles/use policy/CR reports/efficiency goal, lower on governance influence/AI-specific footprint.
9	Intel	Medium	Medium transparency via principles/CR reports/efficiency focus, lower on governance influence/AI-specific footprint.
10	AWS/Amazon	Medium	Medium transparency via RAI guidance/tools/corp reports, lower on governance board details/data center

			specifics/incident responses.
11	Vector Institute	Medium	Medium transparency via public principles, research focus.
12	Amii	Medium	Medium transparency via public principles/framework, research/application focus.
13	Mila	Medium	Medium transparency via partner rules, research ethics focus.
14	BAAI	Medium	Medium transparency via published Beijing Principles.
15	Meta	Medium	Medium transparency via corp reports/some tools, lower on centralized governance/platform harm accountability/AI environmental specifics.
16	Cohere	Medium	Medium transparency via Secure AI Framework, lower on governance details/environmental reporting/incident handling.
17	Cerebras Systems	Low	Low transparency overall, focus on alliances/hardware

			efficiency.
18	Mistral AI	Low	Low transparency overall, limited public docs on governance/environment/bias.
19	Huawei	Low	Low transparency on AI principles/governance, corp report exists, faced external scrutiny.
20	OVHCloud	Low	Low transparency on AI specifics, focus on corp governance/data center efficiency.
21	Alibaba Cloud	Low	Low transparency on AI principles/governance/bias incidents, corp info available.
22	Graphcore	Low	Low transparency overall, supplier ethics main public doc.
23	Groq	Low	Low transparency overall, minimal public information.
24	DeepSeek	Low	Low transparency overall, minimal public information.
25	OpenRouter	Low	Low transparency overall, aggregator role limits scope.

Table 7: Overall Ranking Based on Transparency and Accountability (Weighted)

Rank	Entity	Weighted Score (Approx)	Key Strengths (Transparency & Accountability Focus)	Key Weaknesses (Transparency & Accountability Focus)
1	Microsoft	2.75	Detailed governance, principles, tools; Acknowledges AI env. impact; Proactive commitments (copyright, supplier CFE); RAI reports.	Scope 3 rising, past bias issues took time to address fully.
2	Google/Alphabet	2.70	Detailed principles, reports, governance; Specific TPU LCA; Acknowledged Gemini bias publicly.	Dataset transparency limited; Overall emissions/water use rising linked to AI.
3	Anthropic	2.65	Detailed RSP/Constitutional AI; High governance/risk assessment transparency; Proactive stance on safety/alignment.	Lacks public environmental report/data; Fewer public bias tools.
4	IBM	2.45	Clear principles/pillars; Public Ethics	Lacks AI-specific environmental

			Board details; Fairness tools available; Sunset facial recognition.	reporting; Lower public profile on recent incident handling.
5	OpenAI	1.95	Increased governance transparency (Safety Committee); Unique structure public.	Lacks environmental report/data; Limited bias tool/data transparency; Faces ongoing lawsuits/scrutiny .
6	AWS/Amazon	1.60	RAI guidance/tools; Corp sustainability report; Met renewable goal.	Lacks AI-specific env. data; Governance board less clear; Criticized for data center/labor transparency; Past bias issues.
7	AMD	1.55	Principles/Use Policy; Specific AI efficiency goal; Supplier env. goals; Corp report.	Limited governance process details; Lacks AI-specific env. data (beyond product); Limited incident response record.
8	AI2	1.50	High transparency via open model/data initiatives; AI for Env focus.	Lacks operational environmental report; Research focus limits scope of governance/acc

				ountability data.
9	NVIDIA	1.45	Principles; Corp report detailing efficiency; Powers green systems.	Limited governance process details; Lacks AI-specific env. data (beyond product); Limited incident response record; Criticized over China sales tactics.
10	Intel	1.45	Principles; Corp report detailing efficiency; Human Rights policy link.	Limited governance process details; Lacks AI-specific env. data (beyond product); Limited incident response record.
11	Vector Institute	1.40	Clear principles; Convening role on Green AI/Safety.	Lacks operational environmental report; Research focus limits scope.
11	Amii	1.40	Clear principles/framework; AI for sustainability focus.	Lacks operational environmental report; Research/application focus limits scope.
13	Mila	1.40	Partner rules/enforcement; Research	Lacks operational environmental

			ethics focus; Efficient cluster.	report; Research focus limits scope.
14	Cohere	1.35	Secure AI Framework; Discusses practical risks.	Limited governance details; Lacks environmental report; Limited incident response record.
15	Meta	1.30	Corp sustainability report; Some fairness tools/initiatives; Acknowledges platform risks (human rights policy).	Lacks AI-specific env. data; Governance less clear; Significant criticism on platform harms/bias/accountability.
16	BAAI	1.30	Published Beijing Principles.	Lacks operational environmental report; Limited info on governance implementation/ accountability.
17	Cerebras Systems	1.05	Promotes efficiency; Green AI Cloud partnership; Alliance participation.	Minimal governance transparency; Lacks environmental report; Faced G42 scrutiny.
18	Huawei	1.00	Corp sustainability report; Discusses governance	Lacks clear AI principles; Limited governance transparency;

			ideas; iCooling tool.	Faced lobbying scrutiny.
19	Mistral AI	1.00	Open source focus implies some transparency.	Minimal governance/environmental transparency; Faces data privacy complaint.
20	OVHCloud	1.00	Corp governance detailed; Discusses data center efficiency.	Lacks AI-specific principles/reporting/actions.
21	Alibaba Cloud	0.95	Corp info available; Offers Energy Expert tool; Founder statements on RAI.	Lacks official AI principles/governance details; Limited bias transparency; Lacks environmental report.
22	Graphcore	0.85	Supplier ethics policy.	Minimal AI governance/environmental transparency.
23	Groq	0.85	Promotes LPU efficiency.	Minimal AI governance/environmental transparency.
24	DeepSeek	0.85	External mentions of potential efficiency.	Minimal AI governance/environmental transparency; Concerns over data/guardrails.

25	OpenRouter	0.85	Aggregator role.	Minimal transparency relevant to own operations/governance.

(Note: Weighted scores are approximate, calculated based on assigning High=3, Medium=2, Low=1 to the dimensional scores and applying the weights from Section II. They serve to illustrate the relative ranking based on the framework.)

Industry-Wide Patterns and Insights:

- **Varied State of Transparency:** The analysis confirms a highly uneven landscape regarding transparency and accountability in the AI sector. While some leading companies provide substantial documentation, many players, especially newer or smaller ones, offer minimal public insight into their ethical governance or environmental impact.
- **Common Strengths and Weaknesses:** Most entities publicly articulate high-level ethical principles (e.g., fairness, safety). However, common weaknesses include a lack of transparency regarding the actual effectiveness and influence of internal governance structures, insufficient detail on training data for major models, and, most notably, a significant deficit in specific, verifiable reporting on AI's environmental footprint.
- **Divergence by Company Type:**
  - *Large Tech/Cloud Giants (Google, Microsoft, IBM, AWS, Meta):* Tend to have the most comprehensive (though still imperfect) public reporting across ethics and environment, likely driven by public scrutiny, regulatory anticipation, and resources. However, they also face the most significant documented ethical controversies and environmental impacts due to their scale.
  - *Specialized AI Labs (OpenAI, Anthropic, Cohere, Mistral):* Show varying levels of transparency. Anthropic and OpenAI provide significant detail on safety/governance approaches, while others offer less. Environmental reporting is generally absent.
  - *Hardware Companies (NVIDIA, AMD, Intel, Cerebras, Graphcore, Groq):* Transparency often focuses on product performance and energy efficiency. Disclosure on internal AI ethics governance and operational/lifecycle environmental impact (beyond product use) is typically less detailed.
  - *Research Institutes (AI2, Mila, Vector, Amii, BAAI):* Prioritize openness in research principles and collaborations. Operational environmental reporting

and accountability for deployed systems (if applicable) are less emphasized in public materials.

- **Influence of External Scrutiny:** Public disclosures and accountability measures often appear reactive, driven by media reporting, NGO campaigns (e.g., Greenpeace on data centers [142], Amnesty Tech on surveillance/human rights [145], AI Now on climate/ethics [138]), academic research (e.g., bias studies [74]), employee activism [138], and legal/regulatory pressure.[79] This underscores the vital role of external actors in pushing for greater transparency.
- **The Scope 3 and Absolute Impact Challenge:** The analysis highlights a critical gap in environmental accountability. While operational efficiency (Scope 1 & 2) is often addressed, the rapidly growing Scope 3 emissions (from hardware manufacturing, supply chains, data center construction) and the sheer increase in absolute energy/water demand due to AI scaling are frequently underreported and inadequately addressed by current mitigation strategies and commitments [Insight 9].
- **Persistent Principles-Practice Gap:** Despite widespread adoption of AI ethics principles and the availability of fairness tools from some vendors, high-profile incidents of bias and other ethical failures continue to occur [Insight 2, Insight 3, Insight 5]. This suggests ongoing challenges in effectively embedding ethical considerations throughout the AI development lifecycle and ensuring accountability mechanisms have real teeth.

V. Critical Considerations & Limitations

The analysis and rankings presented in this report must be interpreted within the context of significant limitations inherent in assessing AI ethics and sustainability based solely on publicly available information. These caveats are crucial for a responsible understanding of the findings:

- **Reliance on Publicly Available, Often Self-Reported Data:** The assessment is fundamentally constrained by the information companies choose to disclose publicly. This data is often self-reported, potentially curated to present the organization in the most favorable light, and may not fully reflect internal realities, challenges, or negative outcomes. There is an inherent risk of "ethics washing" or "greenwashing," where public statements and reports overstate positive actions or obscure negative impacts.[138] Access to internal processes, datasets, audit results, and decision-making logs, which would be necessary for a definitive assessment, is unavailable.
- **Lack of Independent Verification:** Many claims made in company reports and statements, particularly regarding the effectiveness of internal governance processes, the mitigation of bias in specific models, or the precise environmental impact reductions from specific initiatives, lack rigorous, independent, third-party verification. While corporate sustainability reports may receive overall assurance [104], this assurance rarely extends to the specific, granular claims made about AI ethics or AI-related environmental data unless explicitly stated. Google's TPU LCA, for instance, appears to be an internal assessment.[112]
- **Inconsistent Metrics and Reporting Frameworks:** The AI industry currently lacks widely adopted, standardized metrics and reporting frameworks for both ethical performance and AI-specific environmental impacts. Companies use varying definitions for concepts like "fairness," different methodologies for calculating energy efficiency (e.g., comparing GPUs vs. CPUs [128], reporting PUE [23]), and inconsistent approaches to scoping and reporting GHG emissions attributable to AI. This lack of standardization severely hinders direct, apples-to-apples comparisons between entities.
- **Transparency vs. Underlying Performance:** This report ranks entities based on their transparency and accountability as evidenced by public data. A high ranking indicates a greater willingness to disclose information and engage publicly on these issues. It does not necessarily equate to superior underlying ethical or environmental performance. An opaque company might, hypothetically, have excellent internal practices but choose not to disclose them, resulting in a low score here. Conversely, a company with significant challenges might be very transparent about them, potentially scoring higher on transparency metrics

despite poorer performance.

- **Dynamic and Evolving Field:** The field of AI, along with societal understanding of its ethical implications and environmental costs, is evolving at an extremely rapid pace. Technologies, best practices, regulatory landscapes, and company policies are constantly changing. This report represents a snapshot based on information available up to [Current Date]. Practices and disclosures may have changed subsequently, potentially altering the rankings.
- **Geopolitical and Cultural Context:** The entities assessed operate within diverse geopolitical and cultural contexts, particularly comparing Western companies with those based in China (Huawei, Alibaba Cloud, BAAI, DeepSeek). Regulatory requirements, stakeholder expectations, and cultural norms regarding corporate disclosure differ significantly across regions [167], influencing the type and extent of information made public. These contextual factors should be considered when interpreting the rankings.

Implications for Interpretation: Readers should treat the rankings presented as indicators of relative public transparency and demonstrable accountability concerning AI ethics and sustainability, rather than as definitive judgments of absolute performance or moral standing. The report highlights which companies are engaging more openly on these critical issues based on available evidence, but it also underscores the pervasive data gaps that prevent a truly comprehensive assessment.

VI. Conclusion: Advancing Responsible AI Through Transparency

This comparative analysis of 25 key AI players reveals a sector grappling with the profound ethical and environmental implications of its transformative technology. While awareness of responsible AI principles is widespread, the translation of these principles into consistently transparent and accountable practices remains highly uneven across the industry. Leading technology firms like Microsoft, Google, IBM, and Anthropic demonstrate greater willingness to disclose governance structures and engage with ethical challenges publicly, yet even they face scrutiny regarding the effectiveness of these measures and the environmental costs of their expanding AI operations. A significant transparency gap persists, particularly concerning the specific environmental footprint of AI—energy consumption, water usage, and lifecycle emissions—which is rarely disaggregated from overall corporate reporting. Furthermore, the effectiveness of internal governance mechanisms and bias mitigation tools often remains opaque, and Scope 3 emissions associated with the AI value chain represent a growing challenge that current actions may not adequately address.

Meaningful progress toward a future where AI is developed and deployed responsibly necessitates a fundamental shift towards greater transparency and accountability across the entire ecosystem. This requires moving beyond voluntary, often curated disclosures towards more systemic and verifiable approaches. Key areas for advancement include:

- **Standardized Reporting:** The development and widespread adoption of standardized reporting frameworks are crucial. This includes standards for disclosing the results of bias audits, the effectiveness of ethical governance processes, and detailed metrics on AI's environmental impact, such as energy and water consumption per computational task (training and inference) and comprehensive lifecycle assessments (LCAs) for AI hardware, potentially building on efforts like Google's TPU LCA [112] and aligning with existing protocols like the GHG Protocol but applied specifically to AI.[21]
- **Independent Verification:** Establishing credible mechanisms for independent, third-party auditing and verification of company claims related to AI ethics and sustainability is essential to build trust and move beyond self-assessment.[50] This could involve specialized auditors or certification bodies focused on AI responsibility.
- **Increased Regulatory Oversight:** Given the limitations of voluntary initiatives and the potential for significant societal harm, regulatory bodies worldwide are increasingly considering measures to govern AI. Regulations like the EU AI Act [29]

may play a critical role in mandating minimum standards for risk management, transparency (e.g., disclosure of training data, system capabilities), and accountability, potentially compelling greater disclosure across the industry.[2]

In conclusion, while this report offers a valuable comparative snapshot of transparency and accountability based on currently available public information, it simultaneously underscores the urgent need for more robust, standardized, and verifiable disclosures. Achieving the promise of AI that is truly beneficial for humanity and the planet requires a collective commitment from industry, policymakers, and civil society to foster an environment where transparency is the norm, not the exception, and accountability is not just stated, but demonstrably enforced.

Glossary

- **Accountability:** The obligation of AI actors to take responsibility for the proper functioning and outcomes of AI systems, especially when harm occurs.
- **Artificial General Intelligence (AGI):** Hypothetical AI that can understand, learn, and apply intelligence across a wide range of tasks, comparable to human cognitive abilities.
- **Artificial Intelligence (AI):** A broad field encompassing the development of computer systems capable of performing tasks that typically require human intelligence, such as learning, problem-solving, and decision-making.
- **Algorithmic Bias:** Systematic and repeatable errors in AI systems that result in unfair discrimination against certain individuals or groups.
- **Bias:** See Algorithmic Bias
- **Carbon Footprint:** The total greenhouse gas emissions caused by an organization, event, product, or AI system, expressed in terms of carbon dioxide equivalent (CO2e).
- **Explainability:** The degree to which a human can understand the cause of an AI's decision or prediction.
- **Inference:** The operational phase of an AI system, where a trained model is used to generate predictions or decisions based on new data.
- **LCA:** Life Cycle Assessment. A standardized methodology for quantifying the environmental impacts of a product or service throughout its entire lifespan.
- **LLM:** Large Language Model. A type of AI model trained on vast amounts of text data, capable of understanding and generating human-like language.
- **PRA (Probabilistic Risk Assessment):** A systematic framework for quantifying the risk of adverse events, used to evaluate both the likelihood and severity of potential outcomes, even those with low probability but high impact.
- **PUE (Power Usage Effectiveness):** A metric that measures a data center's energy efficiency; a lower PUE indicates greater efficiency.
- **Scope 1 Emissions:** Direct greenhouse gas emissions from sources owned or controlled by the reporting entity.
- **Scope 2 Emissions**: Indirect greenhouse gas emissions from the generation of purchased energy.
- **Scope 3 Emissions:** Indirect emissions that occur in the value chain of the reporting entity, including both upstream and downstream emissions.
- **TDL (Deontic Temporal Logic):** An extension of deontic logic that incorporates temporal operators to specify and verify ethical properties of AI systems as they persist or change over time.
- **Training:** The process of teaching an AI model to learn from data.

- **Transparency:** The clear disclosure of how an AI model was created, the data used, and how it functions, providing stakeholders with a comprehensive understanding of the system.
- **XAI:** Explainable Artificial Intelligence.
- **X-Risk:** Existential risk, referring to events that could cause human extinction or irreversibly curtail humanity's potential.
- **WCAG (Web Content Accessibility Guidelines):** Guidelines for making web content more accessible to people with disabilities.
- **WUE (Water Usage Effectiveness):** A metric that measures how efficiently a data center uses water, calculated as the total water used divided by the energy consumption of the IT equipment.

Works cited

1. Veolia and Mistral AI_ join forces to revolutionize resource efficiency management with generative AI and accelerate the ecological transformation, accessed April 30, 2025, https://www.veolia.com/en/our-media/press-releases/veolia-and-mistral-ai-join-forces-revolutionize-resource-efficiency
2. The Ethical Governance of Artificial Intelligence: Using the AI Integrative Risk-Based (AIRB) Model to Achieve Compliance Risk Management - University of Miami School of Law Institutional Repository, accessed April 30, 2025, https://repository.law.miami.edu/cgi/viewcontent.cgi?article=1454&context=umblr
3. The World According to Generative Artificial Intelligence, accessed April 30, 2025, https://carnegieendowment.org/research/2025/01/the-world-according-to-generative-artificial-intelligence?lang=en
4. AI uses more energy, water than a Google search without AI | wcnc ..., accessed April 30, 2025, https://www.wcnc.com/article/news/verify/ai/ai-electricity-water-use-chatgpt-google-la-fire-fact-check/536-85df815e-4212-4f2d-87bd-f5dd4e8f5eff
5. www.foodandwaterwatch.org, accessed April 30, 2025, https://www.foodandwaterwatch.org/wp-content/uploads/2025/03/FSW_0325_AI_Water_Energy.pdf
6. Mitigating AI's environmental impact: A path to sustainable innovation - Tech For Good, accessed April 30, 2025, https://www.techforgood.net/thoughtleadership/mitigating-ais-environmental-impact-a-path-to-sustainable-innovation
7. Environmental impact of AI - Big AI's dirty secret - IMD Business School, accessed April 30, 2025, https://www.imd.org/news/artificial-intelligence/big-ais-dirty-secret-the-environmental-cost-of-generative-ai/
8. Professor's TED Talk warns of AI's hidden water costs | UCR News | UC Riverside, accessed April 30, 2025, https://news.ucr.edu/articles/2025/03/05/professors-ted-talk-warns-ais-hidden-water-costs
9. Explained: Generative AI's environmental impact | MIT News, accessed April 30, 2025, https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117
10. Solving the AI energy dilemma - WTW, accessed April 30, 2025, https://www.wtwco.com/en-nl/insights/2025/03/solving-the-ai-energy-dilemma
11. AI's Environmental Impact: Calculated and Explained - Arbor.eco, accessed April 30, 2025, https://www.arbor.eco/blog/ai-environmental-impact
12. Is AI's energy use a big problem for climate change? | MIT Climate Portal, accessed April 30, 2025, https://climate.mit.edu/ask-mit/ais-energy-use-big-problem-climate-change
13. Sustainable AI: Reduce Your Carbon Footprint, Cut Costs, and Keep Your Users Happy, accessed April 30, 2025,

https://element84.com/company/sustainable-ai-reduce-your-carbon-footprint-cut-costs-and-keep-your-users-happy/
14. Why AI uses so much energy—and what we can do about it, accessed April 30, 2025, https://iee.psu.edu/news/blog/why-ai-uses-so-much-energy-and-what-we-can-do-about-it
15. What does DeepSeek mean for AI's environmental impact? | illuminem, accessed April 30, 2025, https://illuminem.com/illuminemvoices/what-does-deepseek-mean-for-ais-environmental-impact
16. Solving the AI energy dilemma - WTW, accessed April 30, 2025, https://www.wtwco.com/en-cm/insights/2025/03/solving-the-ai-energy-dilemma
17. Green AI Cloud and Cerebras Systems Bring Industry-Leading AI Performance and Sustainability to Europe, accessed April 30, 2025, https://cerebras.ai/press-release/green-ai-cloud-and-cerebras-systems-bring-industry-leading-ai-performance-and-sustainability-to-europe
18. Saudi Arabia invests $1.5 billion in Groq's AI - Green Prophet, accessed April 30, 2025, https://www.greenprophet.com/2025/02/saudi-arabia-invests-1-5-billion-in-groqs-ai/
19. AI and Sustainability: The Power of Integration | Bain & Company, accessed April 30, 2025, https://www.bain.com/insights/ai-and-sustainability-power-of-integration-ceo-sustainability-guide-2024/
20. OpenAI Sustainability: Environmental Impact & Initiatives, accessed April 30, 2025, https://www.byteplus.com/en/topic/546474
21. Environmental Impact of Generative AI | Stats & Facts for 2025 - The Sustainable Agency, accessed April 30, 2025, https://thesustainableagency.com/blog/environmental-impact-of-generative-ai/
22. Spotlight on AI This Earth Day: 'AI Is Fundamentally Incompatible With Environmental Sustainability' - TechRepublic, accessed April 30, 2025, https://www.techrepublic.com/article/news-ai-sustainability-earth-day-2025/
23. Google 2024 Environmental Report - Smart Energy Decisions, accessed April 30, 2025, https://www.smartenergydecisions.com/upload/research_+_reports/google-2024-environmental-report.pdf
24. Microsoft 2024 Environmental Sustainability Report, accessed April 30, 2025, https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/msc/documents/presentations/CSR/Microsoft-2024-Environmental-Sustainability-Report.pdf
25. The Data Center Balancing Act: Powering Sustainable AI Growth | Brown Advisory, accessed April 30, 2025, https://info.brownadvisory.com/The_Data_Center_Balancing_Act_Powering_Sustainable_AI_Growth.pdf
26. The Data Center Balancing Act: Powering Sustainable AI Growth | Brown

Advisory, accessed April 30, 2025, https://www.brownadvisory.com/us/insights/data-center-balancing-act-powering-sustainable-ai-growth
27. GREENHOUSE GAS EMISSIONS FROM AI - Innovation for Cool Earth Forum, accessed April 30, 2025, https://www.icef.go.jp/wp-content/themes/icef/pdf/2024/roadmap/15_ICEF2.0%20GHG%20Emissions%20from%20AI_stand%20alone.pdf
28. Greening Digital Companies Report 2024 1 - World Benchmarking Alliance, accessed April 30, 2025, https://assets.worldbenchmarkingalliance.org/app/uploads/2024/09/Greening-Digital-Companies-Report-2024.pdf
29. Empowering responsible AI practices | Microsoft AI, accessed April 30, 2025, https://www.microsoft.com/en-us/ai/responsible-ai
30. Responsible AI: Key Principles and Best Practices - Atlassian, accessed April 30, 2025, https://www.atlassian.com/blog/artificial-intelligence/responsible-ai
31. Responsible AI - Intel, accessed April 30, 2025, https://www.intel.com/content/www/us/en/artificial-intelligence/responsible-ai.html
32. AI Ethics | IBM, accessed April 30, 2025, https://www.ibm.com/artificial-intelligence/ai-ethics
33. What is AI Ethics? | IBM, accessed April 30, 2025, https://www.ibm.com/think/topics/ai-ethics
34. Building AI Responsibly - AWS, accessed April 30, 2025, https://aws.amazon.com/ai/responsible-ai/
35. Responsible AI: Why Do We Need It? - Artificial Intelligence Board of America, accessed April 30, 2025, https://www.artiba.org/blog/responsible-ai-why-do-we-need-it
36. Responsible AI Guiding Principles - Booz Allen, accessed April 30, 2025, https://www.boozallen.com/insights/ai-research/responsible-ai-guiding-principles.html
37. Responsible AI - AMD, accessed April 30, 2025, https://www.amd.com/en/solutions/ai/responsible-ai.html
38. RRA's Responsible AI Principles - Russell Reynolds Associates, accessed April 30, 2025, https://www.russellreynolds.com/en/about/who-we-are/responsible-ai-principles
39. The ethics of open and public AI: Balancing transparency and safety - Red Hat, accessed April 30, 2025, https://www.redhat.com/en/blog/ethics-open-and-public-ai-balancing-transparency-and-safety
40. AI and Responsible AI - Artificial Intelligence Zone, accessed April 30, 2025, https://www.artificialintelligencezone.com/ai/responsible-ai/
41. Responsible AI Labs, accessed April 30, 2025, https://responsibleailabs.ai/
42. IBM Calls For Rules to Curb Bias in Artificial Intelligence - Insurance Journal, accessed April 30, 2025, https://www.insurancejournal.com/news/national/2020/01/22/555445.htm

43. Ethical concerns mount as AI takes bigger decision-making role - Harvard Gazette, accessed April 30, 2025, https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/
44. Responsible AI Playbook for Investors - World Economic Forum, accessed April 30, 2025, https://www3.weforum.org/docs/WEF_Responsible_AI_Playbook_for_Investors_2024.pdf
45. Corporate Governance Guidelines - Alphabet Investor Relations, accessed April 30, 2025, http://abc.xyz/investor/board-and-governance/corporate-governance-guidelines
46. Artificial Intelligence overview - Microsoft Service Assurance, accessed April 30, 2025, https://learn.microsoft.com/en-us/compliance/assurance/assurance-artificial-intelligence
47. What is AI Governance? Principles & Examples Explained - University of San Diego Online Degrees, accessed April 30, 2025, https://onlinedegrees.sandiego.edu/ai-governance/
48. Anthropic's Transparency Hub \ Anthropic, accessed April 30, 2025, https://www.anthropic.com/voluntary-commitments
49. AI Governance, Public Policy and Safety | Mila, accessed April 30, 2025, https://mila.quebec/en/ai4humanity/ai-governance-public-policy-and-safety
50. www.ai4europe.eu, accessed April 30, 2025, https://www.ai4europe.eu/sites/default/files/2021-09/Huawei%20Interview%20for%20AI4EU_1.pdf
51. What Is AI Governance? - Palo Alto Networks, accessed April 30, 2025, https://www.paloaltonetworks.com/cyberpedia/ai-governance
52. Structure of Meta's Oversight Board | Download Scientific Diagram - ResearchGate, accessed April 30, 2025, https://www.researchgate.net/figure/Structure-of-Metas-Oversight-Board_fig2_378239309
53. What is AI Governance? - IBM, accessed April 30, 2025, https://www.ibm.com/think/topics/ai-governance
54. Our structure | OpenAI, accessed April 30, 2025, https://openai.com/our-structure/
55. An update on our safety & security practices - OpenAI, accessed April 30, 2025, https://openai.com/index/update-on-safety-and-security-practices/
56. OpenAI forms safety committee as it starts training latest artificial intelligence model, accessed April 30, 2025, https://www.courthousenews.com/openai-forms-safety-committee-as-it-starts-training-latest-artificial-intelligence-model/
57. Privacy Progress | Meta, accessed April 30, 2025, https://www.meta.com/actions/privacy-progress/
58. Responsible AI Principles and Approach | Microsoft AI, accessed April 30, 2025, https://www.microsoft.com/en-us/ai/principles-and-approach

59. Data Governance in the Age of Generative AI | AWS Cloud Enterprise Strategy Blog, accessed April 30, 2025, https://aws.amazon.com/blogs/enterprise-strategy/data-governance-in-the-age-of-generative-ai/
60. Singapore's Approach to AI Governance - PDPC, accessed April 30, 2025, https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework
61. Corporate governance | OVHcloud, accessed April 30, 2025, https://corporate.ovhcloud.com/en/company/governance/
62. Strategic Governance of AI: A Roadmap for the Future, accessed April 30, 2025, https://corpgov.law.harvard.edu/2025/04/24/strategic-governance-of-ai-a-roadmap-for-the-future/
63. AI Board Governance Roadmap | Deloitte US, accessed April 30, 2025, https://www2.deloitte.com/us/en/pages/center-for-board-effectiveness/articles/board-of-directors-governance-framework-artificial-intelligence.html
64. Introducing Cohere's Secure AI Frontier Model Framework, accessed April 30, 2025, https://cohere.com/blog/secure-model-framework
65. A Human Rights-Based Approach to Governance and Management - BSR, accessed April 30, 2025, https://www.bsr.org/files/BSR-A-Human-Rights-Based-Approach-to-Governance-and-Management.pdf
66. At the Forefront of AI Innovation - Cerebras, accessed April 30, 2025, https://www.cerebras.ai/company
67. What's Groq AI and Everything About LPU [2025] - Voiceflow, accessed April 30, 2025, https://www.voiceflow.com/blog/groq
68. Ai2 Provides the First-Ever Look Inside an LLM With the Launch of OLMoTrace | Morningstar, accessed April 30, 2025, https://www.morningstar.com/news/business-wire/20250409191408/ai2-provides-the-first-ever-look-inside-an-llm-with-the-launch-of-olmotrace
69. ABOUT THE VECTOR INSTITUTE OUR VISION OUR MISSION GOVERNANCE - AWS, accessed April 30, 2025, https://s3.ca-central-1.amazonaws.com/vectorinstitute.ai/resources/Vector_Backgrounder.pdf
70. About - Alberta Machine Intelligence Institute (Amii), accessed April 30, 2025, https://www.amii.ca/about
71. BAAI/IndustryCorpus_finance · Datasets at Hugging Face, accessed April 30, 2025, https://huggingface.co/datasets/BAAI/IndustryCorpus_finance
72. Best Free AI Models You Can Use on OpenRouter - Apidog, accessed April 30, 2025, https://apidog.com/blog/free-ai-models/
73. Why Google's AI tool was slammed for showing images of people of ..., accessed April 30, 2025, https://www.aljazeera.com/news/2024/3/9/why-google-gemini-wont-show-you-white-people
74. Dr. Joy Buolamwini on Algorithmic Bias and AI Justice | Sanford ..., accessed April 30, 2025,

https://sanford.duke.edu/story/dr-joy-buolamwini-algorithmic-bias-and-ai-justice/
75. cryptorank.io, accessed April 30, 2025, https://cryptorank.io/news/feed/324a8-nvidia-faces-concerns-over-ai-bias#:~:text=Former%20Nvidia%20employees%2C%20Masheika%20Allgood.meeting%20with%20CEO%20Jensen%20Huang.
76. Claude's Defiance: The End of Human Control Over AI? - The Geopolitics, accessed April 30, 2025, https://thegeopolitics.com/claudes-defiance-the-end-of-human-control-over-ai/
77. The EU's Enduring Ethical Deficit in the Aftermath of Huawei - Verfassungsblog, accessed April 30, 2025, https://verfassungsblog.de/ethical-deficit-huawei/
78. Meta Stumbles with AI-Generated Social Media Accounts: What ..., accessed April 30, 2025, https://opentools.ai/news/meta-stumbles-with-ai-generated-social-media-accounts-what-went-wrong
79. OpenAI's Ghibli-style tool raises privacy and data issues | Digital Watch Observatory, accessed April 30, 2025, https://dig.watch/updates/openais-ghibli-style-tool-raises-privacy-and-data-issues
80. OpenAI lawsuit reignites privacy debate over data scraping - CyberScoop, accessed April 30, 2025, https://cyberscoop.com/openai-lawsuit-privacy-data-scraping/
81. OpenAI Faces Lawsuit Alleging Misuse of Internet Users' Data - Investopedia, accessed April 30, 2025, https://www.investopedia.com/openai-faces-lawsuit-alleging-misuse-of-internet-users-data-7555183
82. Addressing AI Bias and Fairness: Challenges, Implications, and Strategies for Ethical AI, accessed April 30, 2025, https://smartdev.com/addressing-ai-bias-and-fairness-challenges-implications-and-strategies-for-ethical-ai/
83. Anti-Jewish and Anti-Israel Bias Found in Leading AI Models, New ADL Report Finds, accessed April 30, 2025, https://www.adl.org/resources/press-release/anti-jewish-and-anti-israel-bias-found-leading-ai-models-new-adl-report
84. The ethical dilemmas of AI | USC Annenberg School for Communication and Journalism, accessed April 30, 2025, https://annenberg.usc.edu/research/center-public-relations/usc-annenberg-relevance-report/ethical-dilemmas-ai
85. How We're Getting AI Risk Wrong - Cohere, accessed April 30, 2025, https://cohere.com/blog/how-were-getting-ai-risk-wrong-aidan-gomez
86. Mistral AI Faces Data Privacy Controversy Over Chatbot 'Le Chat' - OECD.AI, accessed April 30, 2025, https://oecd.ai/en/incidents/2025-02-11-f208
87. Cerebras backer G42 accused of aiding Chinese interests - The Register, accessed April 30, 2025, https://www.theregister.com/2023/11/28/cerebras_g42_china_refile/

88. The Intersection of Innovation and Ethics: Prof. Karina Vold and U of T Panel of Experts Debate AI Safety | Institute for the History & Philosophy of Science & Technology - University of Toronto, accessed April 30, 2025, https://ihpst.utoronto.ca/news/intersection-innovation-and-ethics-prof-karina-vold-and-u-t-panel-experts-debate-ai-safety
89. ByteDance's Doubao LLM Sparks Controversy on Boox E-Readers with Pro-China Bias | AI News - OpenTools, accessed April 30, 2025, https://opentools.ai/news/bytedances-doubao-llm-sparks-controversy-on-boox-e-readers-with-pro-china-bias
90. mila.quebec, accessed April 30, 2025, https://mila.quebec/sites/default/files/media-library/pdf/10589/mila-general-rules-of-conduct-for-partnersfinaldocx.pdf
91. The Algorithmic Tightrope and the Perils of Big Tech's Dominance in AI, accessed April 30, 2025, https://www.technewsworld.com/story/the-algorithmic-tightrope-and-the-perils-of-big-techs-dominance-in-ai-179708.html
92. Responsible AI: Our 2024 report and ongoing work - Google Blog, accessed April 30, 2025, https://blog.google/technology/ai/responsible-ai-2024-report-ongoing-work/
93. Overreliance on AI: Risk Identification and Mitigation Framework - Learn Microsoft, accessed April 30, 2025, https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/overreliance-on-ai/overreliance-on-ai
94. Module 3: AI Bias Mitigation | OLCreate - The Open University, accessed April 30, 2025, https://www.open.edu/openlearncreate/mod/book/tool/print/index.php?id=227374
95. AI Fairness: A Deep Dive Into Microsoft's Fairlearn Toolkit - OpenSense Labs, accessed April 30, 2025, https://opensenselabs.com/blog/ai-fairness
96. How Meta is working to assess fairness in relation to race in the U.S. across its products and systems - Regulations.gov, accessed April 30, 2025, https://downloads.regulations.gov/NTIA-2023-0001-0052/attachment_2.pdf
97. Microsoft Responsible Ai Framework Overview | Restackio, accessed April 30, 2025, https://www.restack.io/p/responsible-ai-practices-answer-microsoft-framework-cat-ai
98. Gender Neutralisation for Unbiased Speech Synthesising | AVESİS, accessed April 30, 2025, https://avesis.yildiz.edu.tr/yayin/37105efb-6f4c-4ef1-b635-e86cda00e76a/gender-neutralisation-for-unbiased-speech-synthesising/document.pdf
99. AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias | Request PDF - ResearchGate, accessed April 30, 2025, https://www.researchgate.net/publication/328091822_AI_Fairness_360_An_Extensible_Toolkit_for_Detecting_Understanding_and_Mitigating_Unwanted_Algorithmic_Bias

100. Manish Raghavan* & Pauline T. Kim** - Georgetown Law Technology Review, accessed April 30, 2025, https://georgetownlawtechreview.org/wp-content/uploads/2024/01/Raghavan_Kim_Final-Proof.pdf
101. Limitations of the “Four-Fifths Rule” and Statistical Parity Tests for Measuring Fairness - Washington University Open Scholarship, accessed April 30, 2025, https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1449&context=law_scholarship
102. ai.google, accessed April 30, 2025, https://ai.google/static/documents/ai-responsibility-update-published-february-2025.pdf
103. 2024 Environmental Report - Google Sustainability, accessed April 30, 2025, https://sustainability.google/reports/google-2024-environmental-report/
104. Sustainability Addendum to Huawei 2024 Annual Report, accessed April 30, 2025, https://www-file.huawei.com/-/media/corp2020/pdf/sustainability/huawei_sustainability_addendum_en_2024.pdf?la=en
105. SSE Jin Qiao Data Center: Using iCooling@AI to Explore Green Finance, accessed April 30, 2025, https://digitalpower.huawei.com/my/success-stories/616.html
106. ESG - Investor Relations :: Intel Corporation (INTC), accessed April 30, 2025, https://www.intc.com/esg
107. 2024 Sustainability Report - Meta Sustainability, accessed April 30, 2025, https://sustainability.atmeta.com/2024-sustainability-report/
108. Think | IBM | State of Sustainability Readiness Report, accessed April 30, 2025, https://www.ibm.com/think/reports/sustainability-readiness
109. Tech vs Carbon: How IBM Achieves Sustainability Goals | Technology Magazine, accessed April 30, 2025, https://technologymagazine.com/articles/tech-vs-carbon-how-ibm-achieves-sustainability-goals
110. Training and using ChatGPT uses a lot of energy, but exact numbers are tricky to pin down without data from OpenAI - Science Feedback, accessed April 30, 2025, https://science.feedback.org/training-and-using-chatgpt-uses-a-lot-of-energy-but-exact-numbers-are-tricky-to-pin-down-without-data-from-openai/
111. AI for Climate, Energy, and Ecosystem Resilience at NVIDIA GTC 2025, accessed April 30, 2025, https://developer.nvidia.com/blog/ai-for-climate-energy-and-ecosystem-resilience-at-nvidia-gtc-2025/
112. Life-Cycle Emissions of AI Hardware: A Cradle-To-Grave Approach and Generational Trends - arXiv, accessed April 30, 2025, https://arxiv.org/html/2502.01671v1
113. NVIDIA Sustainability Report - Impakter Index, accessed April 30, 2025, https://impakter.com/index/nvidia-sustainability-report/
114. 2021 NVIDIA CORPORATE SOCIAL RESPONSIBILITY REPORT, accessed April

30, 2025, https://images.nvidia.com/aem-dam/Solutions/documents/FY2021-NVIDIA-CSR-Social-Responsibility.pdf
115. NVIDIA 2018 Sustainability Report - Responsibility Reports, accessed April 30, 2025, https://www.responsibilityreports.com/HostedData/ResponsibilityReportArchive/n/NASDAQ_NVDA_2018.pdf
116. TPUs improved carbon-efficiency of AI workloads by 3x | Google Cloud Blog, accessed April 30, 2025, https://cloud.google.com/blog/topics/sustainability/tpus-improved-carbon-efficiency-of-ai-workloads-by-3x
117. How moving onto the AWS cloud reduces carbon emissions - Amazon Sustainability, accessed April 30, 2025, https://sustainability.aboutamazon.com/carbon-reduction-aws.pdf
118. 2023 Amazon Sustainability Report, accessed April 30, 2025, https://sustainability.aboutamazon.com/2023-sustainability-report.pdf
119. How AI is helping Amazon buildings conserve water and improve energy efficiency around the world, accessed April 30, 2025, https://www.aboutamazon.com/news/sustainability/amazon-ai-buildings-water-energy-efficiency
120. Alibaba Cloud for Sustainability, accessed April 30, 2025, https://www.alibabacloud.com/en/solutions/sustainability?_p_lc=1
121. Greening intelligence: Charting the future of sustainable AI, accessed April 30, 2025, https://assets.ctfassets.net/9crgcb5vlu43/3lK06heBt0eWEEtAR2klre/f4af108b383b2ea42864af08877fadd8/Greening_intelligence_Charting_the_future_of_sustainable_AI_Economist_Impact_Supported_by_Delta.pdf
122. Ørsted sustainability reporting, accessed April 30, 2025, https://orsted.com/en/who-we-are/sustainability/sustainability-report
123. Sustainability in Cloud Services with OVHcloud's Grégory Lebourg - Solidigm, accessed April 30, 2025, https://www.solidigm.com/products/technology/ovhcloud-sustainability-with-solidigm-ssds.html
124. AMD 2023-24 Corporate Responsibility Report: Advancing Sustainability, Collaboration and Inclusion - Investor Relations, accessed April 30, 2025, https://ir.amd.com/news-events/press-releases/detail/1213/amd-2023-24-corporate-responsibility-report-advancing-sustainability-collaboration-and-inclusion
125. 2023-24 corporate responsibility report | amd, accessed April 30, 2025, https://www.amd.com/content/dam/amd/en/documents/corporate/cr/corporate-responsibility-report.pdf
126. AI Energy Efficiency: Reducing Power Consumption in AI Models - Focalx, accessed April 30, 2025, https://focalx.ai/ai/ai-energy-efficiency-reducing-power-consumption-in-ai-models/
127. Secure AI for Energy and Utilities | Cohere, accessed April 30, 2025,

https://cohere.com/solutions/energy-and-utilities
128. images.nvidia.com, accessed April 30, 2025, https://images.nvidia.com/aem-dam/Solutions/documents/FY2023-NVIDIA-Corporate-Responsibility-Report-1.pdf
129. Nvidia AI Chips: Blackwell Ultra GPU Shipments, Energy Efficiency, Data Center Sustainability and Net Zero Reality Checks - Sustainable Tech Partner for Green IT Service Providers, accessed April 30, 2025, https://sustainabletechpartner.com/vertical-market/technology/nvidia-ai-chips-energy-efficiency-data-center-sustainability-and-net-zero-reality-checks/
130. AI for the environment | Ai2, accessed April 30, 2025, https://allenai.org/ai-for-the-environment
131. Green AI for A Greener Planet Expert Roundtable - Vector Institute for Artificial Intelligence, accessed April 30, 2025, https://vectorinstitute.ai/event/green-ai-for-a-greener-planet-expert-roundtable/?wg-choose-original=true
132. AI-Driven Process Optimization: Enhancing Efficiency and Sustainability, accessed April 30, 2025, https://www.amii.ca/use-cases/engineering-industrial-process-optimization
133. For Google's sustainability reporting, AI is a 'game changer' - Trellis Group, accessed April 30, 2025, https://trellis.net/article/google-ai-sustainability-report-writing/
134. Our 2024 Environmental Sustainability Report - Microsoft On the Issues, accessed April 30, 2025, https://blogs.microsoft.com/on-the-issues/2024/05/15/microsoft-environmental-sustainability-report-2024/
135. Google's Environmental Report 2024 - Smart Energy Decisions, accessed April 30, 2025, https://www.smartenergydecisions.com/white-papers/googles-environmental-report-2024/
136. Sustainability Reports & Case Studies, accessed April 30, 2025, https://sustainability.google/reports/
137. accessed December 31, 1969, https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf
138. The Climate Costs of Big Tech - AI Now Institute, accessed April 30, 2025, https://ainowinstitute.org/publications/climate
139. How Does Graphcore Work? – CanvasBusinessModel.com, accessed April 30, 2025, https://canvasbusinessmodel.com/blogs/how-it-works/graphcore-how-it-works
140. Climate Justice and Labor Rights | Part I: AI Supply Chains and Workflows - AI Now Institute, accessed April 30, 2025, https://ainowinstitute.org/publications/climate-justice-and-labor-rights-part-i-ai-supply-chains-and-workflows
141. 2019 Report - AI Now Institute, accessed April 30, 2025, https://ainowinstitute.org/wp-content/uploads/2023/04/AI_Now_2019_Report.pdf

142. Microsoft, Google, Amazon – Who's the Biggest Climate Hypocrite ..., accessed April 30, 2025, https://www.greenpeace.org/usa/microsoft-google-amazon-energy-oil-ai-climate-hypocrite/
143. Greenpeace Finds Amazon Breaking Commitment to Power Cloud with 100% Renewable Energy, accessed April 30, 2025, https://www.greenpeace.org/usa/greenpeace-finds-amazon-breaking-commitment-to-power-cloud-with-100-renewable-energy/
144. Greenpeace blasts Amazon in latest cloud energy report, lauds Apple - GeekWire, accessed April 30, 2025, https://www.geekwire.com/2015/greenpeace-blasts-amazon-in-latest-cloud-energy-report-lauds-apple/
145. Israel's shocking AI tools & Google's complicity in Gaza | The Listening Post - YouTube, accessed April 30, 2025, https://www.youtube.com/watch?v=cYQcT2Lv-y4
146. Technology - Amnesty International, accessed April 30, 2025, https://www.amnesty.org/en/what-we-do/technology/
147. Amnesty Tech - Amnesty International, accessed April 30, 2025, https://www.amnesty.org/en/tech/
148. Technology, Health, and Human Rights: A Cautionary Tale for the Post-Pandemic World, accessed April 30, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC7762924/
149. Cerebras Pioneers Ethical AI Development through Collaborative AI Initiatives, accessed April 30, 2025, https://cerebras.ai/blog/cerebras-drives-collaborative-ai-initiatives-for-an-ethically-grounded-future
150. Anthropic's AI Revolution: Safety, Scalability, and the $7 Billion Path to Ethical AI Dominance - Astra Global, accessed April 30, 2025, https://www.astraglobal.org/tpost/zrp65fyfx1-anthropics-ai-revolution-safety-scalabil
151. www.nvidia.com, accessed April 30, 2025, https://www.nvidia.com/en-us/ai-data-science/trustworthy-ai/#:~:text=We%20believe%20AI%20should%20respect,avoid%20unwanted%20biases%20and%20discrimination.
152. Trustworthy AI For A Better World - NVIDIA, accessed April 30, 2025, https://www.nvidia.com/en-us/ai-data-science/trustworthy-ai/
153. NVIDIA Launches Cosmos World Foundation Model Platform to Accelerate Physical AI Development, accessed April 30, 2025, https://nvidianews.nvidia.com/news/nvidia-launches-cosmos-world-foundation-model-platform-to-accelerate-physical-ai-development
154. AMD Responsible AI Use Policy, accessed April 30, 2025, https://www.amd.com/en/legal/responsible-ai-use-policy.html
155. Supplier Ethics and Compliance - Graphcore, accessed April 30, 2025, https://www.graphcore.ai/supplier-ethics
156. RESPECT AI | Language for good: Embedding responsible AI into NLP - RBC

Borealis, accessed April 30, 2025, https://rbcborealis.com/news/language-good-embedding-responsibility-nlp/
157. 11 Microsoft, accessed April 30, 2025, https://www.aph.gov.au/DocumentStore.ashx?id=5c57b266-dae2-4c7f-a62e-1287b375aca2&subId=757055
158. Global Trends in AI Governance - World Bank Documents and Reports, accessed April 30, 2025, https://documents1.worldbank.org/curated/en/099120224205026271/pdf/P1786161ad76ca0ae1ba3b1558ca4ff88ba.pdf
159. Mistral 7B: The Next-Generation AI for Advanced Automation and Problem-Solving - Zignuts, accessed April 30, 2025, https://www.zignuts.com/ai/mistral-7b
160. Market Scan Report - By Infosys Responsible AI Office, accessed April 30, 2025, https://www.infosys.com/services/data-ai-topaz/insights/responsible-ai.pdf
161. DeepSeek: Legal Considerations for Enterprise Users | Insights | Ropes & Gray LLP, accessed April 30, 2025, https://www.ropesgray.com/en/insights/alerts/2025/01/deepseek-legal-considerations-for-enterprise-users
162. ETHICAL Principles AI Framework for Higher Education | CSU AI Commons, accessed April 30, 2025, https://genai.calstate.edu/communities/faculty/ethical-and-responsible-use-ai/ethical-principles-ai-framework-higher-education
163. Unlock the Power of Generative AI with Alibaba Cloud Model Studio, accessed April 30, 2025, https://www.alibabacloud.com/blog/unlock-the-power-of-generative-ai-with-alibaba-cloud-model-studio_600957
164. Alibaba founder Jack Ma calls for responsible AI amid its wide adoption across China, accessed April 30, 2025, https://www.thestar.com.my/tech/tech-news/2025/04/11/alibaba-founder-jack-ma-calls-for-responsible-ai-amid-its-wide-adoption-across-china
165. Advancing Safe, Trustworthy and Ethical AI for Climate Action - Arup, accessed April 30, 2025, https://www.arup.com/globalassets/downloads/insights/a/ai-for-sustainable-development-in-asean/ai-for-sustainable-development-in-asean.pdf
166. Alibaba Cloud Model Studio, accessed April 30, 2025, https://www.alibabacloud.com/en/product/modelstudio?_p_lc=1
167. National AI Policy/Strategy of India and China: A Comparative Analysis - RIS, accessed April 30, 2025, https://www.ris.org.in/sites/default/files/Publication/DP%20265%20Amit%20Kumar_0.pdf
168. AI GovernAnce ethIcs - the Globethics Library!, accessed April 30, 2025, https://repository.globethics.net/bitstream/20.500.12424/4318987/6/GE_Global_22_isbn9782889316090.pdf
169. Draft Principles - AI Ethics, accessed April 30, 2025, https://ai-ethics.com/past-principles/

170. AI Trust and Safety - Vector Institute for Artificial Intelligence, accessed April 30, 2025, https://vectorinstitute.ai/ai-trust-and-safety/
171. Principled AI - Alberta Machine Intelligence Institute (Amii), accessed April 30, 2025, https://www.amii.ca/about/principled-ai
172. Beijing Artificial Intelligence Principles – International Research ..., accessed April 30, 2025, https://ai-ethics-and-governance.institute/beijing-artificial-intelligence-principles/
173. Welcome to the AEQUITAS Framework for AI Fairness - Read the Docs, accessed April 30, 2025, https://aequitas-home.readthedocs.io/en/latest/
174. Mila Launched Canada's First AI Computing Cluster Dedicated to Academic Research, accessed April 30, 2025, https://mila.quebec/en/news/mila-launched-canadas-first-ai-computing-cluster-dedicated-to-academic-research
175. Customer Carbon Footprint Tool - AWS, accessed April 30, 2025, https://aws.amazon.com/aws-cost-management/aws-customer-carbon-footprint-tool/
176. Energy Solutions | Google Cloud, accessed April 30, 2025, https://cloud.google.com/solutions/energy
177. Cohere Health's Guiding Principles of AI, accessed April 30, 2025, https://coherehealth.com/guiding-principles-of-ai/
178. Anthropic: Claude 3 - Stanford CRFM, accessed April 30, 2025, https://crfm.stanford.edu/fmti/May-2024/company-reports/Anthropic_Claude%203.html
179. AI Watch: Global regulatory tracker - United States | White & Case LLP, accessed April 30, 2025, https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-united-states