The Dual Mandate: Aligning AI Governance with Corporate Sustainability Targets

Published by: Cetus AI

Subtitle: How Enterprise AI Governance Simultaneously Reduces Costs and Carbon Emissions

Date: June 2026

Reference: CETUS-WP-005 | Version 1.0

4.32g CO_2e	30–60%	FY2026	19+ Models
Per frontier AI query (Scope 3 Cat. 1)	Emissions reduction via intelligent routing	AASB S2 mandatory climate disclosure begins	Routed by Songlines Control®

Executive Summary

As enterprise adoption of generative Artificial Intelligence (AI) accelerates, a critical but often overlooked challenge has emerged: the environmental cost of AI inference. For large organisations, unmanaged employee use of AI tools represents a rapidly growing source of Scope 3 Category 1 greenhouse gas emissions. With mandatory climate disclosure regimes such as Australia's AASB S2 coming into effect from FY2026, the inability to measure, report, and reduce AI-related emissions is no longer just an operational blind spot — it is a compliance and sustainability risk.

This white paper examines the intersection of AI governance and corporate sustainability. It outlines how the same mechanisms required for financial cost control and data security — specifically, intelligent model routing and semantic caching — can simultaneously reduce an organisation's AI carbon footprint by 30% to 60%. By deploying a unified control plane like Songlines Control®, enterprises can align the Chief Financial Officer's cost mandate with the Chief Sustainability Officer's net-zero targets, turning AI governance into a measurable sustainability lever.

Product Architecture Note: The Three Modes of Songlines Control®

The platform operates as a single unified control plane with three distinct operating modes, reflecting an organisation's AI maturity journey:

Operating Mode	Primary Focus	Key Capabilities
Observe Mode	Visibility & Telemetry	Shadow AI Detection, Model Inventory, Usage Reporting
Enforce Mode	Inline Governance & Routing	Inline PII Redaction, Policy Enforcement, Immutable Audit Trail
Orchestrate Mode	Agentic Workflow Management	State Externalisation, Context Window Management

Part One: The Hidden Emissions Problem

AI Inference as Scope 3 Emissions

Under the widely adopted Greenhouse Gas (GHG) Protocol, an organisation's emissions are categorised into three scopes. While Scope 1 (direct emissions) and Scope 2 (purchased electricity) are generally well-understood, Scope 3 covers indirect emissions across the value chain. Specifically, Scope 3 Category 1 — "Purchased Goods and Services" — includes the emissions generated by third-party cloud providers and API services used by the enterprise.

Every time an employee queries a third-party generative AI model such as GPT-4, Claude, or Gemini, the computational energy required for inference is attributed to the enterprise as a Scope 3 Category 1 emission. Recent benchmarking from Google Cloud indicates that a single text query to a large frontier model consumes approximately 0.24 watt-hours (Wh) of energy. Independent research estimates this translates to approximately 4.32 grams of CO_2 equivalent (CO_2e) per query, depending on the data centre's grid mix.

When scaled across an enterprise of 10,000 employees using AI tools multiple times daily, this accumulation becomes highly material. The International Energy Agency (IEA) projects that global electricity consumption from data centres will double to approximately 945 TWh by 2030, with AI inference representing a growing share of that demand. Research published in 2025 estimates that AI's annual carbon footprint could reach between 32.6 and 79.7 million tonnes of CO_2 by 2025 — comparable to the annual emissions of a mid-sized nation.

> *"The crucial first step toward achieving ambitious sustainability targets is gaining full visibility over the sources of emissions — including those that are new, growing, and currently invisible."*

For most enterprises today, AI-related emissions are precisely that: invisible. They do not appear in energy bills, they are not captured by existing IT asset management tools, and they are not attributed to the teams or workflows generating them. They accumulate silently inside Scope 3 Category 1, growing with every new AI tool adopted and every workflow automated.

The Scale of the Problem

AI Model	Energy per Query	Est. CO_2e per Query	Annual Impact (10,000 users, 10 queries/ day)
GPT-4 / Frontier Model	~2.9 Wh	~4.32g CO_2e	~157 tonnes CO_2e
GPT-4o-mini / Efficient Model	~0.24 Wh	~0.36g CO_2e	~13 tonnes CO_2e
Cached Response (Orchestrate Mode)	0 Wh	0g CO_2e	0 tonnes CO_2e

Sources: Google Cloud AI Inference Environmental Impact Report (August 2025); Smartly.AI Carbon Footprint Analysis (June 2024); IEA Energy and AI Report (2025).

Part Two: The Regulatory Imperative

AASB S2 and Mandatory Climate Disclosure

The regulatory landscape is shifting from voluntary sustainability reporting to mandatory, auditable disclosures. In Australia, the Australian Accounting Standards Board (AASB) S2 standard mandates climate-related financial disclosures for large entities beginning in the FY2026 reporting period. AASB S2 requires organisations to assess, measure, and disclose their climate-related risks and opportunities, including material Scope 3 emissions, in a format consistent with the International Sustainability Standards Board (ISSB) framework.

CONFIDENTIAL · Cetus AI · cetusai.com.au · © 2026 Cetus AI Pty Ltd · All Rights Reserved

For organisations that have adopted AI at scale, this creates a direct compliance obligation. Scope 3 Category 1 emissions from AI API usage must be measured, attributed, and reported. Without a governance layer to track API usage at the token level, enterprises have no verifiable method to quantify their AI-related emissions for these mandatory disclosures. Auditors are increasingly scrutinising Scope 3 Category 1 figures, and the absence of AI-related emissions data is likely to be treated as a material omission.

ISO/IEC 42001 and AI Management Systems

The emerging ISO/IEC 42001 standard for AI Management Systems explicitly encourages organisations to evaluate the energy profile of their AI systems and align AI applications with broader sustainability goals. ISO 42001 provides a framework for responsible AI governance that encompasses not only data privacy and algorithmic fairness, but also the environmental impact of AI deployment. Organisations seeking ISO 42001 certification will need to demonstrate that their AI governance practices include consideration of energy consumption and emissions.

The Convergence of Governance and Sustainability Obligations

Standard / Regulation	Jurisdiction	AI Sustainability Requirement	Effective Date
AASB S2 Climate Disclosures	Australia	Mandatory Scope 3 Category 1 reporting including AI API usage	FY2026
ISO/IEC 42001 AI Management System	Global	Energy profile evaluation and sustainability alignment for AI systems	Active (2023)
EU AI Act	European Union	Environmental impact disclosures for high-impact AI systems	Phased 2024–2026
GHG Protocol Scope 3 Standard	Global	Purchased goods and services (Cat. 1) includes cloud AI inference	Active

Part Three: The Cost-Sustainability Nexus

Why Unmanaged AI Defaults to Maximum Energy

The core challenge of unmanaged enterprise AI is that employees naturally default to the most capable — and therefore most computationally expensive and energy-intensive — models available, regardless of the task's complexity. Using a trillion-parameter frontier model to summarise a short email is the computational equivalent of chartering a commercial aircraft for a grocery run. The task is completed, but at a disproportionate cost in both dollars and carbon.

Because cloud providers price AI API access based on the computational compute required (measured in tokens), there is a direct, linear relationship between the financial cost of an AI query and its environmental impact. A query routed to GPT-4 costs approximately 12 times more per token than the same query routed to GPT-4o-mini, and consumes approximately 12 times more energy. The financial and environmental incentives are perfectly aligned — but only if the routing decision is governed.

> *Unmanaged AI is a sustainability risk. Managed AI is a sustainability lever. The same mechanism that reduces AI spend by 30–60% simultaneously reduces AI-related Scope 3 emissions by 30–60%. Cost governance and sustainability are the same mechanism.*

This relationship creates a unique strategic opportunity. The technical interventions required to govern AI costs — model routing policies, usage attribution, semantic caching — are the exact same interventions required to reduce AI emissions. There is no trade-off between financial performance and environmental performance. They are achieved through a single governance layer.

CONFIDENTIAL · Cetus AI · cetusai.com.au · © 2026 Cetus AI Pty Ltd · All Rights Reserved

Part Four: How Songlines Control® Enables Sustainable AI

Songlines Control® sits as an inline proxy between an organisation's employees (or internal applications) and external AI models. This architecture provides the necessary telemetry and control to address both the measurement and reduction of AI emissions across all three operating modes.

Observe Mode: Verifiable Measurement for Climate Disclosure

You cannot reduce what you cannot measure. Songlines Control® in Observe Mode provides token-level telemetry for every AI interaction across the enterprise. This data forms an immutable audit trail that captures the specific model invoked, the exact number of input and output tokens consumed, the department or user initiating the request, and the timestamp and cloud region of execution.

This granular data allows sustainability teams to translate token volumes into accurate, auditable Scope 3 Category 1 emission estimates, replacing guesswork with hard data for AASB S2 and other mandatory climate disclosures. The immutable nature of the audit trail means that the data is defensible under regulatory scrutiny — a critical requirement as climate disclosure standards move toward third-party assurance.

Enforce Mode: Intelligent Model Routing

Not all tasks require frontier models. Songlines Control® in Enforce Mode utilises intelligent model routing to dynamically direct queries to the most efficient model capable of handling the request. Simple classification, summarisation, or extraction tasks are routed to smaller, highly efficient models (such as Llama 3 8B or GPT-4o-mini), which consume a fraction of the energy of their larger counterparts. Complex reasoning tasks that genuinely require frontier capability are routed accordingly.

By automatically matching the computational weight of the model to the complexity of the prompt, organisations typically observe a 30% to 60% reduction in both token costs and the associated carbon footprint, with no degradation in output quality for the vast majority of enterprise use cases. This routing is enforced by policy — not left to individual employee discretion — ensuring that sustainability outcomes are systematic rather than aspirational.

Orchestrate Mode: Semantic Caching

The most energy-efficient AI query is the one that never reaches the model. Songlines Control® in Orchestrate Mode employs advanced semantic caching at the edge. When an employee asks a question that is semantically identical or highly similar to a previously answered query, the platform serves the response directly from the cache, delivering sub-5-millisecond response times with zero inference compute.

In enterprise environments where multiple employees perform similar research, drafting, or analysis tasks, cache hit rates of 20% to 40% are common. Each cache hit eliminates 100% of the energy consumption that would otherwise be attributed to that interaction. At scale, semantic caching can represent the single largest contributor to an organisation's AI emissions reduction trajectory.

The Dual Mandate in Practice

Governance Action	Financial Outcome	Sustainability Outcome	Mode
Route simple tasks to efficient models	30–60% reduction in token spend	30–60% reduction in Scope 3 Cat. 1 AI emissions	Enforce
Serve repeated queries from cache	Elimination of redundant API costs	100% emissions reduction for cached queries	Orchestrate
Token-level usage attribution	Cost chargeback by team/ workflow	Auditable Scope 3 data for AASB S2 reporting	Observe
Cloud region routing	Latency and cost optimisation	Workloads directed to renewable-powered data centres	Enforce

CONFIDENTIAL · Cetus AI · cetusai.com.au · © 2026 Cetus AI Pty Ltd · All Rights Reserved

Part Five: From Baseline to Reduction Trajectory

The 14-Day Baseline Assessment

The first step in any AI sustainability programme is establishing a verifiable baseline. Songlines Control® can be deployed in Observe Mode in as little as 10 minutes via a simple API proxy configuration, with no changes to existing employee workflows or AI tools. Within 14 days, the platform generates a comprehensive AI usage baseline that includes a complete model inventory, token volume by department and workflow, an initial Scope 3 Category 1 AI emissions estimate, and a cost attribution report suitable for board and sustainability committee review.

This baseline serves a dual purpose: it provides the data required for AASB S2 disclosure, and it identifies the specific routing and caching opportunities that will deliver the greatest emissions reductions in subsequent phases.

A Phased Maturity Model

Phase	Mode	Duration	Outcome
1. Baseline	Observe	14 days	Complete AI inventory, initial Scope 3 emissions estimate, cost attribution report
2. Govern	Observe + Enforce	30–60 days	Model routing policies deployed, 30–60% cost and emissions reduction achieved
3. Optimise	All Modes	Ongoing	Semantic caching active, agentic workflow governance, continuous reduction trajectory

Conclusion

As enterprises prepare for a future defined by both ubiquitous AI integration and stringent climate disclosure mandates, the governance of AI infrastructure must evolve. Treating AI solely as an IT or productivity initiative ignores its material environmental impact — and the regulatory obligations that now accompany it.

The good news is that the path to sustainable AI does not require a separate sustainability programme. It requires the same governance infrastructure that responsible AI deployment demands for financial, security, and compliance reasons. By implementing a unified governance layer like Songlines Control®, organisations can establish a verifiable baseline for their AI emissions, intelligently route workloads to minimise energy consumption, and eliminate redundant compute through caching.

In doing so, enterprises can confidently scale their AI capabilities while actively supporting their net-zero commitments and meeting their regulatory obligations under AASB S2 and emerging global standards. The dual mandate — cost governance and sustainability — is not a trade-off. It is a single mechanism, deployed once, delivering two outcomes simultaneously.

References

[1] Google Cloud, "Measuring the environmental impact of AI inference," August 21, 2025. Available: https://cloud.google.com/blog/products/infrastructure/measuring-the-environmental-impact-of-ai-inference/

[2] Smartly.AI, "What is the CO2 emission per ChatGPT query?," June 7, 2024. Available: https://smartly.ai/blog/the-carbon-footprint-of-chatgpt-how-much-co2-does-a-query-generate

[3] International Energy Agency, "Energy and AI," 2025. Available: https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

[4] M. K. Kim, T. A. Yoo, and J. B. Chung, "Toward sustainable generative AI: A scoping review of carbon footprint and environmental impacts across training and inference stages," arXiv preprint arXiv:2511.17179, 2025.

CONFIDENTIAL · Cetus AI · cetusai.com.au · © 2026 Cetus AI Pty Ltd · All Rights Reserved

[5] Australian Accounting Standards Board, "AASB S2 Climate-related Disclosures," September 2024. Available: https://standards.aasb.gov.au/aasb-s2-sep-2024

[6] International Organization for Standardization, "ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system." Available: https://www.iso.org/standard/81230.html

[7] Greenhouse Gas Protocol, "Corporate Value Chain (Scope 3) Accounting and Reporting Standard." Available: https://ghgprotocol.org/corporate-value-chain-scope-3-standard

[8] N. Jegham et al., "How hungry is AI? Benchmarking energy, water, and carbon footprint of LLM inference," arXiv preprint arXiv:2505.09598, 2025.

CONFIDENTIAL · Cetus AI · cetusai.com.au · © 2026 Cetus AI Pty Ltd · All Rights Reserved