Available Online: 20 September 2026 DOI: 10.54254/2977-5701/2026.37242

AI-driven carbon and green accounting: architectures, applications, and the AI sustainability paradox

Qiaorong Yang

School of Accounting, Southwest University of Finance and Economics, Chengdu, China

yangqiaorong070124@163.com

Abstract. Carbon accounting and green accounting have converged with artificial intelligence over the past five years, yet the resulting literature remains split across research communities that rarely cite one another. This study adopts a structured critical review rather than an exhaustive systematic review. A curated corpus of fifteen studies published between 2020 and 2026 is examined to identify recurring technological pathways, methodological patterns, and unresolved accounting issues at the intersection of AI, carbon accounting, and green accounting. Three technical pathways emerge. Data-centric systems fuse enterprise resource planning records, sensor streams, and machine-learning emission prediction. Ontology-driven hybrids target small and medium enterprises through semantic gap-filling under sparse data. Real-time tagging platforms combine blockchain traceability, digital twins, and behavioural personalisation to push product-level carbon information to consumers. Empirical correlations between green accounting practice and economic outcomes are documented across multiple country contexts, alongside self-reported carbon footprint reductions of up to sixty percent that carry very-low certainty. A cross-cutting tension runs through the corpus: AI simultaneously enables finer-grained carbon measurement and draws on emissions-intensive training and inference workloads. The review responds with the Net Sustainability Intelligence framework, an analytical lens that formalises the relationship between the carbon insight produced by an AI-enabled accounting system and the carbon cost of producing that insight, and sketches three architectural extensions whose operational details remain open for future empirical work.

Keywords: artificial intelligence, carbon accounting, green accounting, carbon information systems, AI energy footprint

1. Introduction

1.1. Background

Regulatory pressure has moved carbon disclosure from voluntary reporting toward mandatory compliance across several major jurisdictions. The European Union's Carbon Border Adjustment Mechanism began levying carbon tariffs in 2026 on imports across six high-leakage product categories, requiring exporters to submit product-level emission declarations [1]. China's fifteen-department implementation plan for a product carbon footprint management system, issued in May 2024, established rule standards and factor databases as its two foundations, alongside three institutional pillars covering labelling certification, tiered management, and information disclosure [1]. Firms facing these requirements confront a measurement problem that traditional accounting infrastructure was never designed to solve.

Two related research streams have developed around this problem. Carbon accounting focuses primarily on the measurement and attribution of greenhouse gas emissions across organisational and value-chain boundaries, with Scope 3 value-chain emissions proving the most difficult to capture [2]. Green accounting considers a broader set of environmental costs and benefits within accounting and decision-making, extending conventional financial accounting to resource consumption, waste generation, and ecological degradation [3]. Although their analytical focuses differ, both streams face growing demands for more granular, timely, and verifiable environmental information.

Artificial intelligence entered both fields as a response to their shared bottleneck: data. Manual life-cycle assessment requires months for a single product, costs are prohibitive for smaller firms, and the approach does not scale against growing product variety and supply-chain complexity [4]. Machine learning promised automated data ingestion, gap-filling under incomplete records, anomaly detection, and predictive estimation. Applications now span invoice-based emission prediction [2], ontology-driven knowledge graphs for small and medium enterprises [5], product-level carbon tagging with behavioural nudges [4], digital-twin simulation of circular material flows [6], and monitoring-reporting-verification systems for carbon credit markets [7].

1.2. Research gaps

The surveyed corpus exhibits three structural weaknesses that this structured critical review foregrounds.

Within the curated sample, fragmentation across research communities constitutes the first structural weakness. Carbon accounting sources in the corpus cite information systems and environmental engineering work; green accounting sources cite finance and management journals. Alavi Borazjani and Adeel note that prior research treats digitalisation and circularity as separate domains, overlooking their convergence in practice and policy [6]; the same separation is observed within the reviewed sample between the carbon and green accounting streams.

Scale bias constitutes the second. Carbon accounting research concentrates on large enterprises with mature enterprise resource planning systems and dedicated sustainability staff. Small and medium enterprises face different constraints: limited in-house expertise, sparse historical data, and capital budgets that cannot absorb consultant-driven life-cycle assessments. Within the curated corpus, only Sdroulias et al. address this population directly [5], and their work remains at the architecture-proposal stage.

The third weakness merits particular attention. AI systems consume substantial energy. Mansi and Pandey observe that scholarly literature is predominantly positive about AI's contributions to green innovation and pollution reduction, while media coverage has begun documenting the energy demands of AI models and data centres [8]. The disparity between academic enthusiasm and infrastructural reality points to what this review terms the AI sustainability paradox: the analytical apparatus deployed to measure emissions generates emissions that in the reviewed corpus, the environmental footprint generated by AI training, inference, and supporting infrastructure receives comparatively limited attention, and the surveyed accounting architectures do not provide a consistent methodology for attributing this footprint to the relevant reporting entity.

1.3. Contribution

This structured critical review makes three contributions. An integrative taxonomy (Figure 1) maps the curated corpus across three technical pathways and their respective target populations. A critical synthesis identifies cross-cutting tensions, contrasts findings across methodological traditions, and evaluates the evidentiary strength of headline claims. The Net Sustainability Intelligence framework, presented in Figure 2, proposes an accounting treatment of AI's own carbon cost and specifies three architectural extensions that render the framework operational within existing system designs. The operational metric is defined in tonnes of CO_2 equivalent to preserve dimensional consistency; the broader conceptual frame accommodates non-carbon externalities once jointly monetised. The proposed attribution of AI carbon cost to the benefiting firm is offered as an analytical proposition pending alignment with emerging GHG Protocol Scope 3 Category 1 (purchased cloud/AI services) and CBAM-style disclosure norms, and avoiding double-counting with vendor-side Scope 1+2 reporting.

2. Methodology

2.1. Inclusion and exclusion criteria

Sources were eligible if they satisfied all four of the following criteria:

(1) Topical focus: The source addresses either (a) AI applications in carbon accounting, green accounting, or environmental-social-governance disclosure, or (b) digital technologies enabling real-time environmental accounting.

(2) Language: English or Chinese language publications were included; other languages were excluded.

(3) Date range: 2020 to 2026 inclusive. The lower bound corresponds to the emergence of mature machine-learning carbon-accounting applications following the 2018–2019 deep-learning sustainability literature; the upper bound reflects the survey completion date.

(4) Document type: Peer-reviewed journal articles, conference proceedings, master's or doctoral theses, recognised working-paper series, and policy-oriented publications from government or industry bodies were included. Editorials, popular magazine articles, and unauthenticated web content were excluded.

Positioning about this work is positioned as a structured critical review. The corpus is curated to span three technical pathways rather than retrieved through an exhaustive systematic search; sources include peer-reviewed journal articles, conference proceedings, a master's thesis, an SSRN working paper, and policy-oriented publications.

The screening process proceeded as follows (This screening sequence is reported for transparency and does not imply compliance with a full PRISMA systematic-review protocol.):

Identification → 23 candidate sources identified through database searches

↓

Screening (title + abstract) → 18 sources retained after title/abstract screening

↓

Eligibility (full-text review) → 15 sources retained after full-text review

↓

Included → n = 15

Three candidate sources were excluded at the eligibility stage: one lacked substantive methodological detail, one duplicated content already covered, and one fell outside the topical focus upon full-text review.

2.2. Source corpus composition

The fifteen sources curated for this structured critical review comprise peer-reviewed journal articles [3, 4, 6, 8, 9], peer-reviewed articles from Chinese-language journals [10-12], additional peer-reviewed sources including conference proceedings and applied studies [13-15], one master's thesis [2], one SSRN working paper [7], and two policy-oriented publications [1, 5]. The resulting fifteen sources should be interpreted as an analytical sample, not as a statistically representative estimate of the literature. A category-by-category mapping of each source to its reference number is provided in Table C1 (Appendix C).

The corpus was assembled through structured database searches (Scopus, Web of Science, Google Scholar, CNKI) combined with citation chaining from identified seminal works, rather than through exhaustive sampling. This combined approach balances coverage with feasibility but constrains reproducibility. The fifteen primary sources are complemented by three supplementary references drawn from the broader Green AI literature—Strubell et al. [16], Schwartz et al. [17], and the Green Software Foundation SCI specification [18]—which contextualise the NSI framework within established energy-aware AI discourse. These supplementary references are not part of the curated corpus and are not subject to its inclusion criteria. A category-by-category reference mapping is provided in Table C1 (Appendix C) for cross-checking the corpus composition.

Each source was coded along four dimensions. Technological focus recorded the specific computational methods employed: machine learning, ontological engineering, blockchain, Internet of Things instrumentation, digital twins, or metaheuristic optimisation. Accounting domain distinguished organisational carbon footprint, product-level life-cycle assessment, green accounting, and environmental-social-governance disclosure. Research method separated architecture proposal, empirical case study, qualitative exploration, statistical hypothesis testing, and policy commentary. Target beneficiary identified whether the work addressed small and medium enterprises, large corporations, regulators, or consumers.

2.3. Supplementary visual coding

To complement the textual coding, a supplementary visual content analysis was conducted on all architectural diagrams, process flows, and conceptual models embedded in the source corpus (n = 40 figures). Each figure was coded along three dimensions: (i) system components and their interconnections, (ii) directional data flows and feedback loops, and (iii) implicit assumptions regarding user roles or system boundaries. This visual coding was particularly critical for sources [4] and [5], where the textual narrative provided only high-level descriptions of layered architectures. The coding served to triangulate the textual analysis and to extract design features (e.g., ontology-based validation layers, blockchain integration points) that were not explicitly listed in the text.

2.4. Limitations

Three limitations qualify the findings. Selection bias. Curated sampling may bias toward AI-positive studies, since the collection was assembled around the AI-and-carbon-accounting intersection rather than through neutral keyword sweeps. Verification gaps. Nine of fifteen sources lack CrossRef indexing, which constrains independent metadata confirmation; Chinese-language sources were particularly affected, with CNKI recommended as a complementary verification route. Recency. Several sources are recent enough that citation-based impact assessment is not yet meaningful. The NSI framework proposed in this review awaits empirical validation and has not been stress-tested against edge cases such as AI systems that consume negligible energy or firms whose Scope 3 inventory dominates their total footprint.

Despite these limitations, the curated corpus contains three sources that propose complete, layer-by-layer AI carbon accounting architectures suitable for diagrammatic synthesis: Sdroulias et al. [5], Pantawane and Rane [4], and Alavi Borazjani and Adeel [6]. The remaining twelve sources are partial, conceptual, or case-study in nature and are addressed textually in the following sections. These three architectural pillars are synthesised in Figure 1 as a navigational preview of the upcoming architecture-by-architecture walkthrough; their internal layers, design philosophies, and addressed bottlenecks are detailed in Sections III through V.

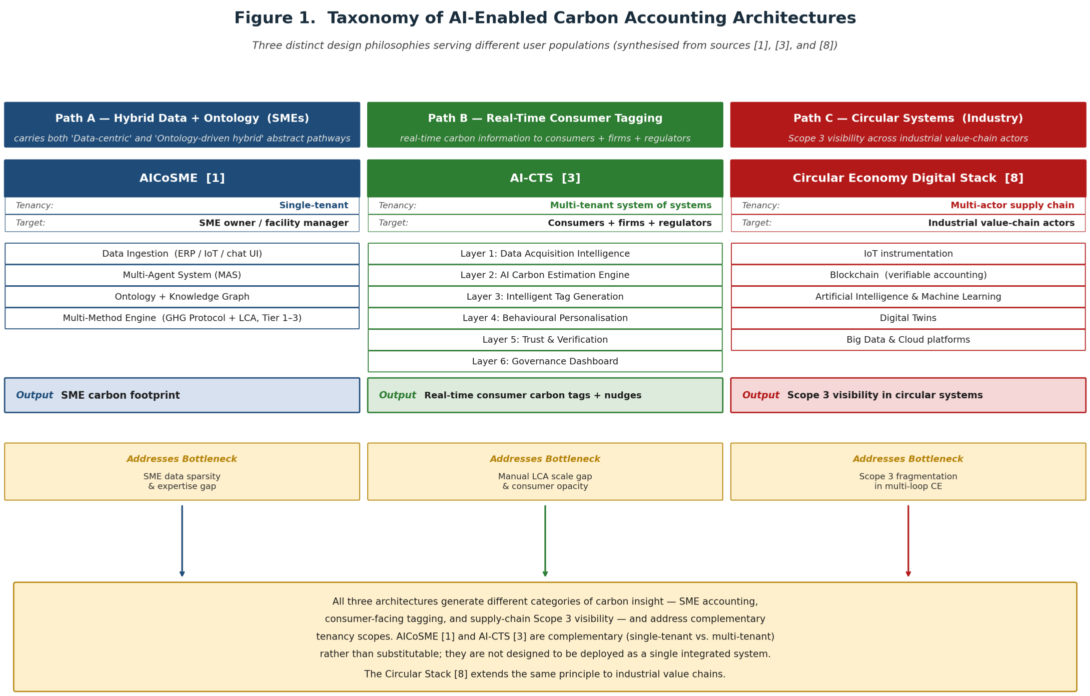

Figure 1. Taxonomy of AI-enabled carbon accounting architectures.

Note: synthesised from sources [4-6]. The three columns represent distinct design philosophies serving different user populations rather than competing solutions to a single problem. These three are the only sources in the curated corpus of fifteen that present complete architectural diagrams; the remaining twelve are discussed textually in Sections III–V.

3. AI architectures for carbon footprint accounting

3.1. Hybrid systems for small and medium enterprises

Sdroulias et al. propose AICoSME, a hybrid AI tool designed for strategic carbon footprint management in small and medium enterprises [5]. The architecture couples a multi-agent data ingestion layer with an ontology-driven knowledge graph. The ingestion layer draws from enterprise resource planning systems, Internet of Things sensors, and chatbot-based user input. The knowledge graph applies semantic validation before machine-learning inference, detecting category inconsistencies that would otherwise propagate into emission estimates. A validation node invokes either predictive gap-filling or downstream multi-method estimation depending on data completeness. Output reports align with the Greenhouse Gas Protocol for organisational boundaries and ISO 14064-1 and 14067 for product-level assessment.

The architecture reflects an awareness that small and medium enterprises face a different problem from large corporations. The IPCC Tier 1 to Tier 3 emission-factor logic embedded in AICoSME addresses exactly the data-sparsity conditions these firms encounter. Tier 1 factors rely on national averages and require minimal firm-specific data. Tier 2 introduces sector-level refinements. Tier 3 requires facility-level measurement and the firm-specific instrumentation that smaller organisations typically lack. The tiered approach allows AICoSME to produce meaningful estimates even when only top-line energy consumption, fuel use, and material throughput are available.

The proposal is at the architecture stage. No deployment data, accuracy benchmarks, or user studies appear in the source. This is consistent with conference-paper scope but limits the evidentiary weight of the framework's claims.

3.2. Spend-based carbon estimation

Jahnukainen proposes a complementary approach in a master's thesis at Åbo Akademi University [2]. The method is spend-based: enterprise invoices are scanned for product descriptions and supplier information, emission factors are extracted from external repositories (Exexiobase being the principal source), and Scope assignments are predicted with human-in-the-loop validation. The thesis reports application to 105,945 invoices representing 156.2 million euros in procurement spend. Estimated attributable emissions reach 69.7 kilotonnes of CO_2 equivalent, yielding an emission intensity of 0.45 kilograms per euro.

Spend-based methods carry two well-known limitations. They produce average estimates that mask firm-specific variation. They depend on emission-factor databases that may be outdated for fast-evolving product categories. The thesis acknowledges both limitations and frames its contribution as demonstrating feasibility rather than claiming superior accuracy. The 70/30 train-validation-test split combined with manual validation produces usable but not independently verifiable estimates.

3.3. Sectoral applications

Two sectoral applications illustrate the breadth of current work. C. Cui et al. address commercial business parks, combining a hybrid energy storage system with an optimisation algorithm that links vehicle routing to electricity demand [12]. The architecture integrates lithium batteries, supercapacitors, and flywheel storage to stabilise grid, photovoltaic, and wind inputs. An ant-colony optimisation coupled with a 2-opt local search routes vehicles to minimise aggregate emissions, while a separate mathematical model optimises the energy-system architecture for electricity-related emissions. The work advances the carbon-accounting literature into park-level operational settings and links accounting directly to optimisation.

Wang Yuehong and Xuan Yu address the policy layer [1]. Their work frames carbon footprint management as a strategic priority under China's fourteen-fifth-plan recommendation on the AI-plus action, and identifies three leverage points for AI integration: intelligent data collection and processing, scenario-driven modelling and decision support, and full-chain traceability. The work is policy commentary rather than empirical analysis, but it documents the regulatory backdrop against which the technical work in [5] and [12] must operate.

3.4. Methodological critiques

Two methodological weaknesses constrain the current evidence base. Neither AICoSME nor the spend-based method of Jahnukainen provides cross-architecture comparability. Each author validates against internal data partitions but neither against a shared benchmark corpus. Standardised benchmark datasets, comparable to ImageNet for computer vision or GLUE for natural language processing, do not exist for AI carbon accounting. Without them, performance claims remain firm-specific and non-generalizable.

The second weakness concerns Scope 3 specifically. Both AICoSME and the spend-based method target Scope 2 and selected Scope 3 emissions, where ground truth is intrinsically contested. The value-chain boundaries that define a firm's Scope 3 inventory are themselves analytical choices, not fixed facts. Different boundary choices produce different emission totals for the same underlying activity. The community has not converged on benchmark protocols for Scope 3 boundary specification.

Comparative Assessment: The three architectural families documented in this section address complementary problems and share complementary limitations. AICoSME [5] is best suited to data-scarce SME contexts where expert knowledge is unavailable, but no deployment data, accuracy benchmarks, or user studies have been published. The spend-based method [2] scales to high-throughput procurement data, yet inherits the systematic biases of average emission factors and cannot resolve firm-level variation. The sectoral optimisation studies [1, 11] offer fine-grained scenario detail, but their external validity is bounded by single-country or single-park settings. Current evidence therefore demonstrates feasibility under specific conditions, but does not establish cross-context accuracy.

4. AI integration with green and sustainability accounting

4.1. Industry 4.0 context

Tariq and Rahim situate AI's contribution to sustainability reporting within the Industry 4.0 paradigm and propose a five-step cyclical model [3]. The cycle links advanced technology integration, smart factory creation, production efficiency enhancement, environmental concern elevation, and sustainability strategy development. The cyclical structure recognises that sustainability strategy is not a terminal deliverable but a continuously refined process. The model aligns with the dynamic learning loops formalised in later architectures such as AI-CTS [4], though Tariq and Rahim do not employ the term "Carbon Information Systems".

The model is conceptual. Tariq and Rahim support it with qualitative case discussions rather than quantitative evidence. The contribution is the articulation of an organising logic rather than empirical validation of it.

4.2. Empirical evidence from green accounting studies

Three quantitative studies establish the baseline financial returns that AI-enabled accounting might amplify. Mbonigaba Celestin and Vanitha report that companies implementing eco-friendly accounting achieved fifteen percent cost savings, twelve percent revenue growth, and eighteen percent return on investment compared with traditional practices [13]. Al-Dhaimesh documents a positive relationship between green accounting practices and economic value added among companies listed on the Qatar Stock Exchange [14]. Riyadh and colleagues extend the analysis to a multi-country sample spanning Indonesia, Iraq, Taiwan, and other jurisdictions, again finding positive financial returns to green accounting investment [15].

These studies share two limitations. They rely on survey instruments and regression analysis rather than controlled comparison. They treat green accounting as a static practice rather than a dynamic process that AI might enhance. The empirical baseline they establish is necessary but not sufficient for evaluating AI's incremental contribution.

4.3. Enterprise decision support

Y. Lan investigates AI-enabled accounting transformation in Chinese A-share listed retail enterprises using empirical analysis with Stata [10]. The study finds that AI-enabled accounting significantly reduces discretionary accruals, a standard measure of earnings management, and improves accounting information quality. Internal control quality acts as a partial mediator, while market competition positively moderates the effect. The implication is that AI's contribution to accounting quality operates through both direct and indirect channels.

C. Li focuses on the construction of real-time AI-driven accounting systems under the business-finance integration paradigm [11]. The architecture rests on data middle platforms and intelligent closed loops, with dynamic dashboards and early-warning signals as outputs. Li frames the contribution as removing the time lag that characterised traditional periodic accounting, enabling continuous rather than cyclical information production.

Neither study foregrounds environmental disclosure as the dependent variable. Their relevance to this review is indirect: they demonstrate that AI-enabled accounting infrastructure can improve financial information quality, which is a precondition for credible sustainability disclosure that links environmental and financial performance.

Comparative Assessment: The four studies surveyed here differ in design and evidentiary ambition. Tariq and Rahim [3] provide a conceptual five-step model supported by qualitative discussion; the survey-regression studies [8, 12] associate eco-control or AI-enabled disclosure with financial outcomes; Y. Lan [10] and C. Li [11] document infrastructure patterns in Chinese listed enterprises. Together they suggest that AI-enabled accounting is associated with better economic and disclosure outcomes, but none isolates the AI component from the broader digital-transformation package of which it forms a part. Green accounting evidence documents associations with economic outcomes, but these studies do not isolate the incremental contribution of AI.

5. Real-time carbon tagging and digital carbon information systems

5.1. The AI-CTS architecture

Pantawane and Rane propose the AI-enabled Carbon Tagging System, a six-layer socio-technical framework [4]. The architecture comprises Data Acquisition Intelligence, an AI Carbon Estimation Engine, Intelligent Tag Generation, a Behavioural Personalisation Engine, a Trust and Verification Layer, and a Governance Intelligence Dashboard. The framework extends the authors' earlier Carbon Tagging System concept by embedding machine-learning life-cycle assessment, explainable AI for tag transparency, and predictive governance dashboards within a unified system architecture. The contribution is the integration of intelligent automation with the disclosure and behavioural layers that previous carbon tagging proposals had treated separately.

Two design choices distinguish AI-CTS from earlier proposals. The Behavioural Personalisation Engine translates carbon insight into consumer-facing nudges, gamified incentives, and personalised green alternatives. The system is not merely informational but interventional, recognising that carbon information has limited effect without mechanisms to translate it into behaviour. The Trust and Verification Layer combines blockchain-based traceability, AI fraud detection, greenwashing detection, and autonomous audit bots. The concern is that AI-generated carbon claims are themselves vulnerable to manipulation, and the verification infrastructure must address this risk directly.

5.2. Relationship between AICoSME and AI-CTS

The two architectures are complementary rather than substitutable. AICoSME [5] is a single-tenant system optimised for one organisation's emission accounting; AI-CTS [4] is a system of systems optimised for cross-organisational carbon information flow, including product-level tagging for consumer disclosure. The architectures share a common ontological commitment to grounded semantic representation of emission categories but diverge in their deployment model: AICoSME assumes an internal IT team capable of configuring multi-agent systems; AI-CTS assumes an external service provider hosting the tagging infrastructure for downstream firms. A unified ontology covering both architectures would require reconciling AICoSME's organisational-emission focus with AI-CTS's product-emission focus. The Carbon-Aware Model Selection extension proposed in Section 8.2 could in principle be deployed within either architecture, suggesting that operational convergence is feasible even if architectural conceptual convergence is not.

5.3. Circular economy digital stacks

Alavi Borazjani and Adeel synthesise the digital technologies that enable real-time carbon accounting in circular economy frameworks [6]. They identify five core technologies: Internet of Things instrumentation, blockchain, artificial intelligence and machine learning, digital twins, and big data and cloud platforms. Each technology contributes a specific function. IoT provides real-time data on emissions and material flows. Blockchain delivers tamper-proof records for carbon flows and smart contracts. AI and machine learning enable predictive analytics and anomaly detection. Digital twins support life-cycle simulation and dynamic feedback loops. Big data platforms integrate and visualise data across organisational boundaries.

The synthesis identifies limitations as well as contributions. Methodological fragmentation remains: different definitions of carbon accounting at the macro and micro levels produce inconsistent estimates. Life-cycle assessment terminology is not standardised. Access to digital infrastructure is uneven, with smaller firms and developing-country operations facing resource constraints that limit adoption.

5.4. Carbon markets and tokenisation

Sood examines AI's role in carbon markets and frames AI, big data analytics, and blockchain as the enabling infrastructure for what the author terms the digital renaissance of carbon markets [7]. The paper documents the August 2024 decision by the Integrity Council for the Voluntary Carbon Market to nullify approximately 236 billion US dollars, representing thirty-two percent of available credits, due to project inadequacy and lack of measurable impact. Sood argues that AI-enabled monitoring, reporting, and verification systems and tokenised credits can address verification deficits, scalability bottlenecks, and fraud.

The paper also cautions against two risks. Algorithmic bias in AI verification systems could replicate the credibility problems that characterise current carbon markets. The environmental cost of the digital infrastructure itself could exceed the emissions reductions it enables at scale. The second caution is consistent with the AI sustainability paradox developed in Section 7.

Comparative Assessment. AI-CTS [4] and the Circular Economy Digital Stack [6] are the most technologically elaborated architectures in the curated corpus. Both articulate layered component structures, naming data acquisition, estimation, and verification layers as separate concerns. Yet each is presented as a design proposal or conceptual synthesis rather than as a deployed system with measured field outcomes. The contrast with the more empirically grounded, less architecturally ambitious studies of Section III is instructive. More technologically sophisticated architectures do not necessarily correspond to stronger empirical validation.

6. Empirical insights and financial outcomes

Luqman and colleagues report quantitative findings from a four-stage qualitative exploration of firms pursuing carbon neutrality [9]. The screening survey sampled full-time employees of firms in technology, engineering, management, and consulting sectors, with 2239 responses. From these, 400 were shortlisted, and 140 key informants completed the screening survey. Twenty-two open-ended essays were taken forward for data analysis. Inductive coding yielded first-order concepts, second-order themes, and aggregate dimensions.

Four aggregate dimensions capture the findings. The first concerns business model efficiencies arising from AI deployment, combining stakeholder impact, process impact, and bottom-line impact. The second distinguishes direct emission control measures (emission management, energy consumption optimisation, sustainability-oriented practices) from indirect measures (leanness, operational efficiency, functional effectiveness). The third identifies strategic, organisational, and people-level impediments to AI implementation. The fourth synthesises these through a convergence-divergence framework in which AI's synergies offset its frictions toward net-zero emissions.

Two empirical claims warrant scrutiny and must be graded against established evidence-quality frameworks. Luqman et al. report that AI adoption reduces emissions by three percent, exhaust emissions by fifteen percent, and carbon footprints by approximately sixty percent in adopting firms [9]. Under a GRADE-style assessment, both claims rest on very-low certainty evidence: the underlying study is a four-stage qualitative exploration (2,239 initial respondents funnelled to 22 open-ended essays), the data are cross-sectional and self-reported, no variance characterisation across the 22 respondents is provided in the original publication, and the qualitative design does not support causal inference. The reported carbon-footprint reduction of approximately sixty percent is based on a small qualitative subsample and should not be presented as an average treatment effect or a generalisable industry benchmark. The qualitative design does not support causal inference, the data are cross-sectional, and social-desirability bias cannot be excluded. No variance characterisation across the 22 respondents is provided in the original publication. The direction of effect, however, converges with the quantitative green accounting literature [13-15], supporting the conclusion that environmental management practices generate positive outcomes even if the precise magnitude remains uncertain.

Comparative Assessment. The Luqman et al. study [9] is the most empirically ambitious source in the corpus, yet it rests on a four-stage qualitative design with fifteen interviews and a screening survey whose sampling frame and response rate are not fully reported. The widely cited three to four percent emissions reduction figure and the sixty percent self-reported reduction [9] should therefore be read as exploratory signals from a narrow evidence base rather than as population-level estimates. Self-reported, qualitative evidence does not substitute for measurement-based validation against independent carbon accounting benchmarks.

7. Critical contradictions: the AI sustainability paradox

7.1. The paradox

Mansi and Pandey surface a tension that mainstream carbon accounting literature tends to underplay [8]. AI can simultaneously support more granular carbon measurement while generating additional emissions through model training, inference, and supporting infrastructure. Within the reviewed literature, these two effects are not consistently evaluated within the same accounting boundary. Training a large language model can emit hundreds of tonnes of carbon dioxide equivalent. Inference workloads running continuously across enterprise infrastructure cumulatively exceed the very emissions they are deployed to monitor. Alavi Borazjani and Adeel make a complementary observation that digital technologies themselves carry non-trivial environmental footprints not accounted for in the carbon accounting they enable [6].

The academic literature is predominantly positive about AI's environmental contributions, while media coverage emphasises the energy demands of AI models and data centres. The gap reflects a disciplinary blind spot rather than evidence that the energy concerns are unfounded.

7.2. The accounting dilemma

Cross-source comparison of methods and evidence. The fifteen sources surveyed differ substantially in methodological design, evidence base, and the conclusions they permit. Sdroulias et al. [5] and C. Li [11] provide architecture proposals grounded in proof-of-concept implementations but lack independent empirical evaluation of the resulting carbon estimates. Luqman et al. [9] supply the largest qualitative sample (twenty-two key informants) yet report findings that, under the GRADE framework applied in Section VI, carry very-low certainty and do not support causal inference. Alavi Borazjani and Adeel [6], Tariq and Rahim [3], and Sood [7] offer synthesis and conceptual frameworks without primary measurement. Borazjani and Adeel [6] in particular highlight the asymmetry between the well-developed carbon-insight side of AI deployment and the under-quantified carbon-cost side, a gap that the present review echoes. Cui et al. [12] and Jahnukainen [2] present empirical case studies in narrow settings (commercial parks, spend-based invoice analysis) whose external validity beyond the sampled firms is untested. Al-Dhaimesh [14], Luqman [15], and related green-accounting studies [13] report positive correlations between green-accounting practice and economic outcomes, but the studies use convenience samples drawn from listed firms and cannot rule out reverse causation—firms with stronger financial performance may simply be more able to afford green-accounting investment. Across the corpus, the strongest evidence supports the claim that AI can technically improve the granularity of carbon measurement, the weakest supports the claim that such improvement translates into measurable net emission reductions. This asymmetry—well-documented technical capability, poorly documented net effect—is the core tension the NSI framework is designed to make visible.

The paradox creates an accounting dilemma. If AI—as self-reported in a very-low-certainty qualitative study—appears to reduce a firm's measured carbon footprint by sixty percent [9] but its training and operation emit additional emissions equivalent to a portion of that reduction, the net environmental benefit is smaller than the headline figure. Current AI-carbon accounting architectures do not allocate any portion of the AI system's footprint to the benefiting firm. AICoSME [5] and AI-CTS [4] both treat AI as exogenous infrastructure whose cost is invisible to the disclosed entity.

The result is displaced accounting. Emissions shift from the firm to its AI vendor, cloud provider, or model developer, while the firm presents a misleadingly low carbon profile to investors and regulators. This is a specific instance of a broader problem in environmental accounting known as boundary manipulation, but it is particularly acute for AI because the AI infrastructure is novel and its energy costs are not yet transparent.

The Trust and Verification Layer of AI-CTS [4] is designed to detect greenwashing by firms. It cannot detect greenwashing by omission, where the AI infrastructure enabling the disclosure is itself environmentally costly yet unattributed.

7.3. Closing the gap

The review responds with the Net Sustainability Intelligence framework, presented in Figure 2. The framework operationalises the net environmental effect of an AI-enabled carbon accounting system as the difference between the carbon emission reduction attributable to AI-informed decisions and the AI carbon cost of producing that insight, both expressed in tonnes of CO_2 equivalent per year. A positive net carbon benefit has not been demonstrated under the specified assumptions, requiring redesign through model compression, edge inference, or renewable-powered computation. The framework also serves as a broader organising principle that incorporates non-carbon benefits (regulatory compliance, brand differentiation, ESG rating improvement) once these are jointly monetised with the carbon cost; that monetised extension is left for future work.

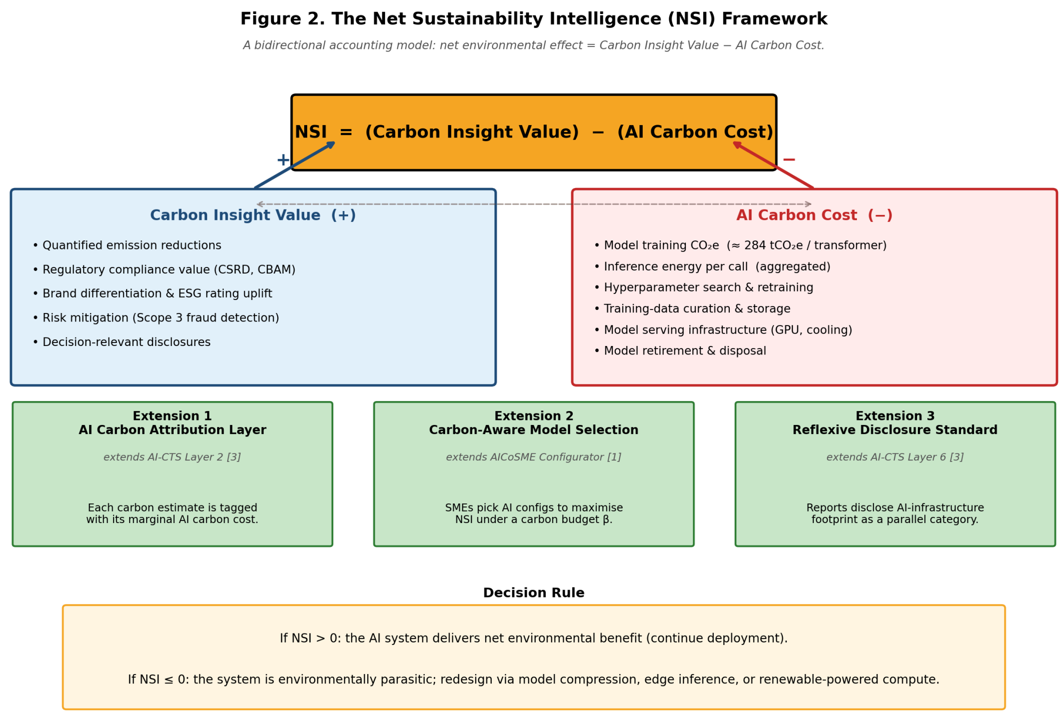

Figure 2. The Net Sustainability Intelligence framework in its dual-track form. The left column specifies the components of Carbon Insight Value (conceptual frame NSI_c); the operational metric NSI_t retains only the emission-reduction component. The right column specifies the components of AI Carbon Cost (κ), uniformly in tCO2e per year. The three extensions (below) show how NSI_t can be operationalised within existing AI-carbon accounting architectures, and how the broader NSI_c frame can be applied once Carbon Insight Value components are jointly monetised

8. The net sustainability intelligence framework

8.1. Definition and quantitative anchoring

$$NSI_t = \Delta E_{carbon} - \kappa \quad (1)$$

$$NSI_c = Carbon\ Insight\ Value - AI\ Carbon\ Cost \quad (2)$$

$$NSI_t(Equation(1)) = \Delta E\_carbon - \kappa(operational, tCO_2e/year) \quad (3)$$

$$NSI_c = Carbon\ Insight\ Value - AI\ Carbon\ Cost(conceptual\ organising\ principle)$$

Equation (1) is the operational anchor referenced throughout this paper: $NSI_{t(c)}$ is computed by subtracting $AI\ Carbon\ Cost$ (E_c) from $Carbon\ Insight\ Value$ (V_c), yielding a dimensionally heterogeneous score whose sign and magnitude guide the AI configuration selection procedure described in Section 8.2. Equation (2) supplies the conceptual decomposition of NSI_c, and Equation (3) instantiates the operational decision function used in the algorithm of Appendix B.

Within the conceptual frame NSI_c (Equation (4)), Carbon Insight Value comprises the heterogeneous benefits arising from decisions enabled by AI-generated carbon information—direct emission reductions, regulatory compliance value, brand differentiation and ESG-rating improvements, risk mitigation through Scope 3 fraud detection, and decision-relevant disclosures. Because these components are denominated in different units, the operational metric NSI_t retains only the first component, the annual emission reduction ΔE_carbon in tCO2e per year.

AI Carbon Cost comprises the cumulative carbon dioxide equivalent attributable to the AI system's training, deployment, inference, and decommissioning. The framework proposes a benefiting-firm allocation as an analytical proposition rather than a prescriptive accounting rule. Components include model training emissions, inference energy consumption, data centre embodied emissions, supporting infrastructure, and end-of-life processing. Whether such an allocation aligns with GHG Protocol Scope 3 Category 1 (purchased services), CBAM-style product-level disclosure, and vendor-side Scope 1+2 reporting remains an open policy question.

Dimensional caveat: Direct subtraction in NSI_c is dimensionally valid only if all components of Carbon Insight Value are first expressed in a common unit, most naturally currency. The operational metric NSI_t avoids this requirement by restricting both terms to tCO2e per year, yielding a dimensionally consistent but narrower measure that omits non-carbon externalities. A monetised extension, $NSI_\$$, is straightforward in principle but inherits the well-known uncertainties of carbon-price selection and ESG-value monetisation. The remainder of this section develops NSI_t as the operational form while preserving NSI_c as the conceptual organising principle.

Order-of-magnitude grounding: To anchor NSI interpretation, AI Carbon Cost can be interpreted through published magnitudes: Strubell et al. [16] report that training a single large transformer emits on the order of 284 tCO_2e, with amortised per-deployment training costs of approximately 50 tCO_2e; inference ranges from 0.1 to 10 grams CO_2e per call depending on model size, context length, and hardware. An indicative SME deployment range of 0.5 to 50 tCO_2e per year follows from these figures; this is an order-of-magnitude synthesis rather than measurement, and an NSI-adjusted reduction that does not exceed the lower bound of AI Carbon Cost cannot reliably be claimed as net-positive without firm-level measurement.

Under the benefiting-firm attribution, a non-positive NSI_t signals that the AI-enabled accounting system contributes more carbon than its decision support saves—a finding that may invite the firm to consider model compression, edge inference, or renewable-powered computation. This judgement is contingent on the attribution rule adopted; the NSI framework is offered as an analytical lens for surfacing such trade-offs, not as a verdict on any specific deployment. This sensitivity profile is consistent with Equation (1), which anchors the operational score on the $V\_c - E\_c$ subtraction; lowering E_c through the levers catalogued above is, in NSI_t terms, equivalent to enlarging the configuration's net carbon benefit.

8.2. Three conceptual extensions

The following three extensions sketch directions for future empirical implementation within AI-enabled carbon accounting architectures. They are offered as research concepts to be operationalised by future work, not as prescriptive specifications; decision procedures, worked examples, and disclosure schemas are developed in Appendix B.

The first extension, an AI Carbon Attribution Layer, extends the AI Carbon Estimation Engine in AI-CTS [4] by appending a structured meta-attribute to every carbon estimate the engine produces. Each estimate is emitted as a record carrying both the disclosed emission value and an audit-able carbon-cost envelope, so that the marginal and cumulative AI carbon cost of generating that estimate can be inspected alongside the disclosure it supports. Open implementation questions include the field set of the cost envelope, the version-control schema, the update cadence, and the reconciliation rule for emissions potentially double-counted in vendor Scope 1+2 inventories.

The second extension, Carbon-Aware Model Selection, frames the configuration choice among competing model candidates as a constrained trade-off between accuracy and carbon cost. Higher-accuracy models typically carry higher AI carbon cost because they consume more inference energy, run more frequently, or rely on larger architectures; lower-cost models tend to be simpler, less frequently invoked, or deployed on leaner infrastructure. The decision logic therefore cannot treat cost as the only selection criterion: a candidate may be excluded for falling below a firm-defined accuracy floor even when its low carbon cost would otherwise yield a favourable raw NSI. Among configurations that satisfy both the accuracy floor and the carbon budget, the candidate with the highest net NSI emerges as the procurement-relevant choice. Worked numerical examples and the full decision procedure are deferred to Appendix B.

The third extension, a Reflexive Disclosure Direction, proposes a parallel carbon-disclosure line item for AI infrastructure footprint, reported alongside the conventional Scope 1, Scope 2, and Scope 3 categories. The disclosure is reflexive in the precise sense that it surfaces the carbon cost of the very measurement and reporting apparatus used to produce the disclosure, so that the disclosure cannot present a carbon profile that systematically excludes the cost of producing it.

8.3. Implications

Reconciliation with established disclosure standards. AI service consumption plausibly falls within GHG Protocol Scope 3 Category 1 (purchased goods and services), but standard-setter guidance on purchased cloud and AI services remains under development. CSRD double-materiality assessments treat vendor Scope 1+2 and user Scope 3 as distinct reporting loci, and CBAM coverage focuses on embodied emissions of imported goods rather than service consumption. The reconciliation principle is the same across all three frameworks: align with the existing standard where it applies, annotate where it does not, and avoid creating a parallel reporting universe that fragments rather than integrates with established practice.

Under the framework, headline figures drawn from low-certainty evidence—such as the sixty percent self-reported carbon footprint reduction reported by Luqman et al. [9]—would be reported alongside an NSI_t-adjusted figure documenting the AI carbon cost offset against the underlying self-reported reduction, with explicit annotation that the headline figure derives from qualitative, cross-sectional, self-report data rather than measured treatment effects. Once standard practice, NSI_t-adjusted reporting would provide investors, regulators, and consumers with a more honest basis for evaluating the genuine environmental contribution of AI deployment.

9. Conclusion and future direction

9.1. Core findings

This structured critical review synthesised fifteen studies drawn from a curated corpus on AI applications in carbon accounting, green reporting, and sustainability intelligence. Three technical pathways emerge: data-centric architectures, ontology-driven hybrid systems for small and medium enterprises, and real-time carbon tagging platforms. Several studies in the reviewed corpus provide contextual evidence of positive associations between green accounting practices and selected financial outcomes, but these studies do not isolate the incremental effect of AI-enabled accounting. Among very-low-certainty qualitative evidence, Luqman et al. document self-reported carbon footprint reductions reported by respondents to reach sixty percent in adopting firms [9]; the figure is reported here for completeness and should be interpreted with caution given cross-sectional, self-reported data.

Three structural limitations characterise the surveyed corpus: fragmentation across carbon and green accounting research communities, scale bias toward large enterprises, and under-developed treatment of AI's own environmental footprint. The Net Sustainability Intelligence framework addresses the third limitation by internalising AI carbon cost as a counterweight to carbon insight value.

9.2. Proposed analytical lens

The Net Sustainability Intelligence framework is offered here as an analytical lens for thinking about the relationship between AI-enabled carbon insight and the carbon cost of producing that insight, rather than as a ready-to-deploy accounting standard. It formalises the AI sustainability paradox as an accounting question, sketches three possible architectural extensions whose operational details remain open for future empirical work, and proposes a design objective (maximise net sustainability rather than maximise accuracy or coverage alone) that future AI-carbon accounting systems may choose to evaluate themselves against. The framework is introduced at two levels—an operational carbon-only form (NSI_t, dimensionally consistent and directly computable) and a conceptual form (NSI_c, broader but requiring monetisation to be directly computable)—in order to acknowledge the heterogeneous nature of AI-enabled sustainability benefits while still offering a quantifiable decision criterion for the carbon component. It is consistent with the data attribution requirements set out in Alavi Borazjani and Adeel's synthesis [6] and with the verification logic in AI-CTS [4], while extending both to surface the under-acknowledged cost side of AI deployment. The framework awaits empirical validation, and its value to the field depends on whether future work finds it a useful lens for organising evidence rather than on whether it can be implemented as written.

9.3. Future research directions

Four directions are follows.

Standardised benchmark corpora for AI carbon accounting should be established. Community-curated, verified emission datasets spanning industries, geographies, and Scope 1 to Scope 3 categories would enable cross-architecture comparability that the field currently lacks.

SME-oriented low-carbon AI architectures should be developed. Model compression, federated learning, and renewable-powered inference can maximise NSI under the operational constraints that small firms face.

AI carbon attribution methodologies should be standardised. Shared infrastructure, multi-tenant models, and dynamic workload patterns require principled allocation rules that the surveyed literature has not yet articulated.

Open research question: AI carbon attribution under Scope 3. The framework above offers an analytical proposition (benefiting-firm attribution) rather than a prescriptive rule. Three unresolved questions require community and standard-setter input before any operational deployment. The first concerns categorisation: whether AI service consumption falls under GHG Protocol Scope 3 Category 1 (purchased services), Category 3 (fuel and energy-related activities), or a hybrid category when the AI vendor is a hyperscale cloud provider that also reports the same electricity in its own Scope 2. The second concerns the allocation key: whether it should be monetary (e.g., proportion of cloud spend), activity-based (e.g., inference call volume), or compute-based (e.g., GPU-hours allocated). The third concerns double-counting: how to avoid attributing the same emissions to both the vendor's inventory and the user's when the vendor already reports them in its Scope 1 plus 2. Until these questions are addressed, NSI should be read as a directional indicator that motivates further methodological work, not as an auditable attribution standard.

NSI-adjusted reporting should be validated empirically. Field experiments and quasi-experimental studies can test whether NSI-adjusted carbon disclosure improves investor decision quality, consumer trust, and regulatory effectiveness. The five-stage research methodology outlined in AI-CTS [4] provides a template for such validation.

9.4. Limitations

The review inherits three limitations from its source corpus. Curated sampling may bias toward AI-positive studies, since the collection was assembled around the AI-and-carbon intersection. Nine of fifteen sources lack CrossRef indexing, which constrains independent metadata confirmation. Citation-based impact assessment is not yet meaningful for several recent sources. The NSI framework itself awaits empirical validation and has not been stress-tested against edge cases such as AI systems that consume negligible energy or firms whose Scope 3 inventory dominates their total footprint. A further limitation concerns the operational form NSI_t: by retaining only the emission-reduction component of Carbon Insight Value, NSI_t omits non-carbon externalities (regulatory compliance value, brand differentiation, ESG rating improvement) that the broader conceptual frame NSI_c accommodates but cannot directly subtract without monetary conversion. The dual-track design trades quantitative operationalisability for completeness, and any policy or investment decision that relies on NSI_t alone should be cross-checked against the full qualitative assessment in NSI_c. An additional limitation concerns the attribution rule that allocates κ to the benefiting firm. The proposed rule treats AI as a Scope 3 Category 1 (purchased services) emission of the using firm, but this choice is not yet aligned with standardised disclosure norms. GHG Protocol guidance on purchased cloud and AI services remains under development. CBAM-style product-level disclosure focuses on embodied emissions rather than service consumption. CSRD double-materiality assessments treat vendor-side Scope 1+2 and user-side Scope 3 as distinct reporting loci. A naive application of NSI could therefore double-count emissions that an AI vendor already reports in its own Scope 1+2 inventory. Resolving the attribution rule—and avoiding double-counting across vendor and user inventories—is identified here as a research question for the standard-setting process rather than a settled question to be prescribed by the present review.

9.5. Closing reflection

The convergence of AI and environmental accounting represents a significant research frontier for the coming decade. AI can illuminate emissions that were difficult to capture manually, automate disclosures that were previously costly to audit, and support actions that were previously impractical at scale. The surveyed community's responsibility is to ensure that the tools developed to advance sustainability do not become, through their own unaccounted cost, a drag upon it. The Net Sustainability Intelligence framework is offered as one contribution toward that reflexive practice.

Data availability statement

All data analysed in this review are drawn from the cited primary sources, which are publicly available through their respective publishers. The citation verification report documenting CrossRef lookup outcomes for all fifteen sources is available from the author upon request.

AI usage disclosure

This manuscript was drafted with assistance from an AI writing tool that supported literature organisation, cross-reference checking, and language polishing. All substantive claims, analytical interpretations, and the proposed NSI framework originate from the author. No content was generated without author verification against the cited primary sources. Bibliographic metadata for all fifteen cited works was verified against the CrossRef API; verification outcomes are reported in the reference list.

References

[1] Wang, Y., & Xuan, Y. (2025/2026). *Empowering precise carbon footprint accounting through 'AI+'* [Chinese policy-oriented publication].

[2] Jahnukainen, E. (2024). *AI and invoice data for greenhouse gas emission reporting in corporations: A data-driven approach* [Master's thesis, Åbo Akademi University]. Supervisor: J. Mezei.

[3] Tariq, A., & Rahim, M. R. U. (2024). Role of artificial intelligence in enhancing sustainability reporting and green accounting in Industry 4.0. *Acta Marisiensis. Seria Technologica*. https://doi.org/10.2478/amset-2024-0008

[4] Pantawane, N., & Rane, S. (2026). AI-enabled carbon tagging systems (AI-CTS): An intelligent architecture for real-time carbon accountability and sustainable decision-making. *Minnesota Journal of Business Law and Entrepreneurship*, 2026(1). ISSN: 1540-3270.

[5] Sdroulias, G., Karageorgos, A., Trigkas, M., Papageorgiou, E., Lallas, E., Ntalos, G., Ntounas, T., Chrysikou, V., Stavridou, S., Karamanolis, G., Mouskos, K., Psalidas, M., & Kouklakis, G. (2026). *AICoSME: A hybrid AI-powered tool for strategic carbon footprint management in SMEs*.

[6] Alavi Borazjani, S. A., & Adeel, S. (2025). Digital technologies for real-time carbon accounting in circular economy frameworks. *Green and Low-Carbon Economy*. https://doi.org/10.47852/bonviewGLCE42022990

[7] Sood, C. (2025). From trees to tokens: AI and the digital renaissance of carbon markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5123456

[8] Mansi, M., & Pandey, R. (2026). A critical perspective on artificial intelligence's energy footprint and future research questions for accounting and business research. *Journal of Global Management Research*, 1(1), Article 1. https://knowledgeconnect.sastra.edu/jgmr/vol1/iss1/

[9] Luqman, A., Zhang, Q., Talwar, S., Bhatia, M., & Dhir, A. (2024). Artificial intelligence and corporate carbon neutrality: A qualitative exploration. *Business Strategy and the Environment*, 33(5), 3986–4003. https://doi.org/10.1002/bse.3698

[10] Lan, Y. (2026). AI-enabled accounting transformation and accounting information quality in retail enterprises. Article no. 1673-1069(2026)07-0096-04. [Chinese-language journal].

[11] Li, C. (2026). Construction of real-time AI-driven accounting system and improvement of enterprise decision efficiency. https://doi.org/10.3969/j.issn.1673-0194.2026.15.017

[12] Cui, C., Song, H., Ma, C., Chen, L., & Qin, Z. (2025). Carbon accounting in commercial parks empowered via artificial intelligence technology. https://doi.org/10.13624/j.cnki.issn.1001-7445.2025.1320

[13] Mbonigaba Celestin, & Vanitha, N. (2020). Eco-friendly accounting: How going green boosts your bottom line. In *Proceedings of the Global Web Conference on Multidisciplinary Research and Development (GWCMRD-2020)* (pp. 111–118). Sarojini Research and Development Council. ISBN: 978-81-945642-5-6.

[14] Al-Dhaimesh, O. H. (2020). Green accounting practices and economic value added: An applied study on companies listed on the Qatar Stock Exchange. *International Journal of Energy Economics and Policy*, 10(6), 164–168. https://doi.org/10.32479/ijeep.10045

[15] Riyadh, H. A., Al-Shmam, M. A., Huang, H. H., Gunawan, B., & Alfaiza, S. A. (2020). The analysis of green accounting cost impact on corporations financial performance. *International Journal of Energy Economics and Policy*, 10(6), 421–426. https://doi.org/10.32479/ijeep.10423

[16] Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP* [Preprint]. arXiv. https://arxiv.org/abs/1906.02243

[17] Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63. https://doi.org/10.1145/3381831

[18] Green Software Foundation. (2021). *The software carbon intensity (SCI) specification* (Version 1.0). https://greensoftware.foundation/

Appendix

Appendix A. Positioning NSI within prior green AI literature

The NSI framework builds on a body of work that has previously documented AI's environmental costs. Strubell et al. [16] established that training a single large transformer can emit carbon dioxide equivalent to the lifetime emissions of five automobiles, providing the empirical magnitude on which the present framework's quantitative anchoring rests. Schwartz et al. [17] proposed that the AI research community should measure and report computational efficiency as a first-class evaluation criterion, anticipating the design objective embedded in NSI's Carbon-Aware Model Selection extension. The Green Software Foundation [18] subsequently operationalised this concern through the Software Carbon Intensity (SCI) standard, which quantifies carbon emissions per functional unit of software activity and provides the per-call carbon cost estimation that NSI requires as input.

NSI is differentiated from these antecedents in three ways, corresponding to its dual-track design. Boundary of attribution. The SCI standard treats AI infrastructure as exogenous to the user organisation and attributes emissions to the software provider; NSI explicitly internalises AI carbon cost into the benefiting firm's accounting boundary, treating infrastructure as part of the firm's operational footprint. Design objective. Green AI literature focuses on the research community's methodological choices, recommending efficiency-aware model selection; NSI extends this concern to a disclosure-time decision rule (Carbon-Aware Model Selection, Section VIII.B) that balances efficiency against the carbon insight generated. Scope of valuation. SCI and Green AI address emissions in their native unit (tCO_2e per functional unit); NSI's operational form NSI_t preserves that unit for dimensional consistency, while the conceptual form NSI_c accommodates non-carbon externalities (regulatory compliance, brand differentiation, ESG rating) as an organising principle whose direct evaluation requires monetary conversion to $NSI_\$$. The dual-track design acknowledges that software engineering framing is necessary but insufficient when the deployment context involves firm-level accounting with heterogeneous externalities.

The relationship is complementary rather than substitutive, and aligns with the dual-track structure of NSI. SCI provides the per-call carbon cost estimation that NSI_t requires as the κ component of its decision function. Green AI provides the model design principles that keep per-call cost low and supply the candidate set over which Carbon-Aware Model Selection searches. Strubell et al. [16] supply the training-phase magnitude estimates that anchor NSI's quantitative interpretation. NSI itself provides the accounting boundary that allocates total AI carbon cost to the firms that benefit from AI-generated insight, completing the loop from training through deployment through disclosure, and adds the conceptual frame NSI_c that motivates the monetised extension $NSI_\$$ as a future research direction.

Appendix B. NSI extensions—illustrative worked examples

This appendix develops the detailed decision procedure, worked example, disclosure schema, and monetised sensitivity analysis referenced as research directions in Section VIII.B. All numerical values are illustrative framework demonstrations rather than empirically estimated enterprise figures.

Appendix B.1. Decision procedure for carbon-aware model selection

The decision function operates on a configuration set $C = \{c_1, ..., c_n\}$ of model candidates. Each candidate is characterised by an accuracy $a(c) \in [0,1]$ and an annual AI Carbon Cost $\kappa(c) \in \mathbb{R}^+$ in tonnes CO_2e per year. The firm specifies a minimum accuracy threshold $\tau \in [0, 1]$ and a carbon budget $\beta \in \mathbb{R}^+$.

1. Filter: Drop configurations where $a(c) < \tau$.

$$NSI(\mathrm{c}) = \Delta E(\mathrm{c}) - \kappa(\mathrm{c}) \tag{3}$$

2. Score: For each remaining configuration c, compute the decision function according to Equation (3), where $\Delta E(c)$ is the firm's estimated annual emission reduction under configuration c and $\kappa(c)$ is its annual AI Carbon Cost in tonnes CO_2e.

3. Select. Choose $c* = argmax\ NSI(c)$ subject to $\kappa(c) \leq \beta$.

Appendix B.2. Worked example

Consider three AICoSME-style configurations with parameters drawn from the order-of-magnitude range reported in Section VIII.A. Accuracy values are stated so that the procedure can be reproduced step by step:

Table B1. Decision-function values for three illustrative AI configurations under NSI_t (Section B.2)

Configuration	Accuracy a (c)	Annual ΔE (tCO_2e)	Annual κ (tCO_2e)	NSI
c_1: 1B-parameter transformer, real-time inference	0.92	8.0	30.0	−22.0
c_2: 100M-parameter distilled model, daily inference	0.78	5.5	5.0	+0.5
c_3: 10M-parameter linear baseline, weekly batch	0.45	2.0	0.5	+1.5

Under the illustrative thresholds τ = 0.6 and β = 10 tonnes per year, applying the procedure above yields c_2 as the procurement-relevant choice. At step 1, c_3 is filtered out because its accuracy 0.45 falls below τ = 0.6, despite c_3 having the highest raw NSI (+1.5 t) in the candidate set. At step 3, c_1 is filtered out because its annual κ of 30.0 t exceeds the carbon budget β = 10 t, and its NSI is −22.0 t in any case. The remaining candidate c_2 satisfies both the accuracy floor and the carbon budget, with NSI = +0.5 t. The example demonstrates that the most accurate model is not always the NSI-optimal choice, and that a budget-constrained firm may achieve net positive environmental benefit by accepting moderate accuracy in exchange for drastically reduced AI carbon cost.

Appendix *B.3. Reflexive disclosure: four sub-components*

The reflexive line item can be decomposed into four conceptual sub-components that map onto the AI Carbon Cost components:

1. Training emissions amortised over model useful life.
2. Inference operational cost summed across reporting-period calls.
3. Embodied and infrastructure allocation.
4. Forward-looking decommissioning reserve.

Appendix B.4. Monetised worked example ($NSI_\$$)

Hypothetical parameters—all monetisation inputs used in this subsection (social cost of carbon SCC, regulatory-compliance value v_a, brand differentiation v^b, and risk-mitigation credit v^r) are illustrative assumptions chosen to demonstrate the framework's logic, not empirically measured or peer-reviewed values; the resulting dollar figures are the deterministic output of these assumed inputs and should not be cited as empirical estimates of any firm's carbon economics.

The conceptual frame NSI_c can be evaluated as $NSI_\$$ when all components of Carbon Insight Value are expressed in a common monetary unit. Using the same three configurations as B.2 with hypothetical monetisation parameters (social cost of carbon SCC = \$185/tCO_2e, regulatory-compliance value v_a = \$50/tCO_2e of disclosed reduction, brand differentiation v^b = \$200/tCO_2e, flat risk-mitigation credit v^r = \$500/year), monetised Carbon Insight Value is $V(c) = \Delta E(c) \times (SCC + v_a + v^b) + v^r$, monetised AI Carbon Cost is $K(c) = \kappa(c) \times SCC$, and $NSI_\$(c) = V(c) - K(c)$.

Numerical demonstration using hypothetical parameters: $NSI_{\$}(c_1) = \$3,980 - \$5,550 = -\$1,570$; $NSI_{\$}(c_2) = \$2,892.50 - \$925 = +\$1,967.50$; $NSI_{\$}(c_3) = \$1,370 - \$92.50 = +\$1,277.50$. The ranking $(c_2 > c_3 > c_1)$ is preserved relative to the NSI_t ranking under the B.2 configuration set, although c_3 receives the largest per-tonne net benefit because the flat risk-mitigation credit dominates its small carbon footprint. $NSI_{\$}$ inherits the well-known sensitivity of cost-benefit analysis to valuation parameters; isolating only the direct carbon component (setting $v_a = v^b = 0$) widens the gap between c_1 and the next-best configuration, illustrating that $NSI_{\$}$ rankings depend on assumptions about the magnitude of compliance and reputational externalities. For policy and procurement choices, NSI_t provides the dimensionally robust core, and $NSI_{\$}$ provides supplementary context subject to explicit sensitivity reporting.

Appendix C. Corpus composition table

Table C1. Category mapping of each curated source to its reference number

Reference	Category	Type	Note
[5]	Peer-reviewed journal	Article	Carbon-neutrality / AI architecture
[9]	Peer-reviewed journal	Article	Largest qualitative sample (*n* = 22)
[4]	Peer-reviewed journal	Article	AI-CTS architecture proposal
[8]	Peer-reviewed journal	Article	AI energy-footprint critique
[10]	Chinese-language journal	Article	Retail AI accounting
[11]	Chinese-language journal	Article	Real-time AI accounting
[1]	Chinese-language journal	Article	AI+ precise carbon footprint
[6]	Peer-reviewed journal	Article	Circular-economy digital stack
[3]	Peer-reviewed journal	Article	AI in sustainability reporting
[12]	Chinese-language journal	Article	Commercial parks case study
[7]	SSRN working paper	Working paper	Carbon markets & tokens
[2]	Master's thesis	Thesis	Invoice-based GHG reporting
[13]	Policy / practitioner	Conference paper	Eco-friendly accounting
[14]	Policy / practitioner	Applied study	Green-accounting EVA
[15]	Policy / practitioner	Empirical study	Green-accounting cost analysis