Creative Economy
2026, VOL. 10, NO. 3, 98-101
DOI: 10.47297/wspceWSP2516-251920.20261003

Artificial Intelligence in ESG Disclosure and Sustainable Finance: From Narrative Reporting to Data-Driven Transparency

Zhi Zheng, Guoqing Chen

School of Finance, Chengdu Jincheng College, Chengdu, Sichuan, 611731, China

ABSTRACT
The shift to mandatory, data-driven ESG disclosure under the CSRD and ISSB standards has generated urgent demand for artificial intelligence. This paper provides a conceptual analysis of AI's reconfiguration of ESG disclosure across four dimensions: NLP-based text mining, machine learning for ESG scoring, AI-enabled assurance, and the regulatory and organizational conditions shaping adoption. We argue that AI transforms ESG reporting into a verifiable, transparent system, but effectiveness hinges on resolving tensions: data noise and standardization gaps, algorithmic greenwashing, ESG rating divergence, and the primacy of organizational AI capability over tool procurement. We propose a research agenda prioritizing multilingual NLP, causal identification, AI governance in assurance, emerging-market adaptation, and the normative implications of delegating credibility to algorithms. The paper integrates fragmented disciplinary perspectives into a unified framework for scholars, regulators, and practitioners.

KEYWORDS
ESG disclosure; Artificial intelligence; Sustainable finance; Natural language processing; Machine learning; Greenwashing detection

1 Introduction

Over the past decade, corporate sustainability reporting has undergone a paradigmatic shift. The European Union's Corporate Sustainability Reporting Directive (CSRD), adopted in 2022 and effective from January 2024, expands mandatory ESG reporting to approximately 50,000 companies under the European Sustainability Reporting Standards, subjecting reports to compulsory third-party assurance[1]. In parallel, the International Sustainability Standards Board issued IFRS S1 and IFRS S2 in June 2023, establishing a globally consistent baseline adopted by more than 20 jurisdictions[2]. These milestones mark an inflection point: ESG reporting is now a core governance mechanism with direct implications for capital allocation and firm valuation. The scale and complexity of these obligations have outpaced conventional manual reporting workflows. A multinational firm subject to CSRD must compile, validate, and report hundreds of ESG data points across supply chains and jurisdictions, generating demand for technological solutions capable of automating data capture and supporting independent verification[3]. Artificial intelligence has emerged as the central enabler, offering tools ranging from natural language processing for textual analysis to machine learning for ESG scoring, generative AI for report drafting, and satellite imagery for real-time environmental verification[4]. Despite the rapid proliferation of AI applications in sustainable finance, the academic literature remains fragmented across disciplinary silos. Computer science contributions often emphasize algorithmic performance without engaging institutional contexts; accounting and finance scholarship frequently treats AI as a black-box tool rather than as a technology reshaping the epistemological foundations of sustainability disclosure itself . This paper addresses this gap by providing a conceptual analysis of how AI is transforming ESG disclosure from a narrative, self-reported practice into a data-driven, verifiable system. The analysis is organized around four dimensions: NLP and text mining for disclosure analysis; machine learning for ESG scoring and corporate performance; AI-enabled assurance and auditing; and regulatory and organizational conditions enabling or constraining AI adoption. The paper contributes an integrated conceptual framework, identifies critical tensions, and proposes a structured future research agenda.

2 The Regulatory Transformation and Rise of AI

The origins of modern corporate sustainability disclosure trace to voluntary CSR reporting in the 1990s and early 2000s, with the social foundations of ESG emerging from early rating agencies and institutional investors . These early disclosures were characterized by narrative richness, qualitative orientation, and lack of standardization. They functioned primarily as legitimacy-seeking devices, signaling commitment without necessarily providing decision-useful, comparable, or verifiable data [5]. The limitations of this voluntary regime became apparent as sustainable finance matured: investors faced a data environment characterized by incompleteness, inconsistency, and low comparability, providing the institutional impetus for mandatory reporting. The contemporary regulatory landscape is defined by three overlapping frameworks. The Task Force on Climate-related Financial Disclosures, established by the Financial Stability

Board in 2017 and now largely superseded by the ISSB, provided the first widely adopted framework for climate-related financial disclosures. By 2023, the ISSB had absorbed the TCFD framework into IFRS S2, establishing a formal global standard[2]. The CSRD mandates that large companies and listed SMEs report comprehensive sustainability information under the European Sustainability Reporting Standards, introducing requirements that fundamentally alter ESG disclosure: the principle of double materiality; mandatory inclusion of sustainability reports within the management report; compulsory third-party assurance; and digital, machine-readable tagging[1]. The ISSB standards adopt a complementary approach, focusing on sustainability-related financial information material to investors' assessment of enterprise value, with jurisdictions including the United Kingdom, Australia, Singapore, and Nigeria having committed to adoption or endorsement. The convergence of CSRD and ISSB frameworks signals a trajectory toward mandatory, standardized, and globally aligned sustainability disclosure [6,10]. This regulatory expansion has generated an operational imperative for digital transformation. Under the CSRD, a reporting entity must manage hundreds of data points spanning environmental, social, and governance metrics. The double materiality assessment adds complexity, as firms must evaluate both their own sustainability impacts and how external risks translate into financial risks. The digital tagging requirement, aligned with the European Single Electronic Format, necessitates structured data infrastructure. Within this context, AI has emerged as a transformative force. The field of AI applications in ESG disclosure has grown at a compound annual growth rate of 91.9% between 2020 and 2025, reflecting the convergence of large language model maturation and intensifying regulatory pressure [4]. Generative AI is anticipated to reconfigure the reporting value chain, reducing compliance costs while increasing the volume and granularity of disclosed information[3].

3 AI Technologies in ESG Disclosure

The largest and most mature stream of AI applications centers on natural language processing and text mining. Sustainability reports, CSR disclosures, and integrated reports are predominantly textual documents historically assessed through manual content analysis—a method that is labor-intensive, subject to coder bias, and poorly scalable. NLP offers a systematic alternative, enabling automated extraction, classification, and evaluation of textual features at scale. Foundational work established the viability of corpus-based NLP analysis for measuring the readability of sustainability reports, demonstrating that standard readability formulae and NLP-based techniques produce quantifiable, reproducible metrics of textual complexity[7]. The field has since expanded to encompass sentiment analysis, topic modeling, greenwashing detection, and materiality identification. BERT-based models applied to climate disclosure texts in the European energy sector illustrate how domain-adapted NLP tools capture sector-specific nuances that generic language models miss [8]. Greenwashing detection constitutes a particularly active substream. Environmental claim detection models using transformer architectures identify discrepancies between stated commitments and measurable outcomes, assessing the specificity, consistency, and verifiability of sustainability claims[9]. The convergence of textual analysis with external performance metrics represents a critical frontier for regulatory technology, particularly under CSRD third-party verification requirements[10]. Large language models analyzing the relationship between manager sentiment, policy uncertainty, and ESG disclosure quality demonstrate that AI-powered textual analysis can surface predictive signals invisible to conventional quantitative methods[11]. The second major dimension encompasses supervised and unsupervised machine learning algorithms to predict ESG scores, assess disclosure quality, and examine the relationship between ESG performance and corporate financial outcomes. Studies employ gradient boosting machines, XGBoost, K-means clustering, and interpretability tools such as SHAP values to enhance predictive accuracy and transparency. A pivotal contribution analyzing approximately 400 companies in the Euro-Stoxx 600 Index over 2011–2020 establishes that firm-level ESG scores are significantly associated with financial performance across multiple sectors[12]. Machine learning techniques have also identified institutional governance predictors of disclosure behavior—such as board secretaries' competence, salary, functional experience, age, and tenure—extending the analytical scope beyond conventional firm financials[13]. Unsupervised machine learning reveals latent disclosure profiles at the organizational level, identifying distinct archetypes across industry sectors invisible to traditional statistical approaches[14]. Cross-sector machine learning applied to carbon-intensive industries demonstrates that algorithmic analysis surfaces sector-specific performance patterns with material implications for climate risk assessment[15]. The third dimension addresses the transformation of independent verification from a document-centric, retrospective process into a data-driven, continuous, and forward-looking assurance system. The Audit 4.0 framework provides the conceptual foundation, prescribing real-time assurance by linking the physical and digital worlds using sensors, Internet of Things devices, and satellite imagery[16]. In ESG assurance, this paradigm enables auditors to obtain exogenous evidence from the physical world rather than relying solely on management-provided documents. A proof-of-concept using satellite images to estimate methane concentrations from oil and gas facilities demonstrates how satellite-based GHG emission verification can be integrated into ESG audit evidence, cross-referencing corporate disclosures against satellite-derived data to detect discrepancies invisible under conventional document-review assurance[17]. Continuous auditing technologies monitor ESG performance indicators on an ongoing basis rather than through periodic reviews[18]. The International Auditing and Assurance Standards Board's development of ISSA 5000 explicitly contemplates the role of technology in assurance evidence gathering, suggesting that regulatory frameworks are beginning to institutionalize the AI-assurance convergence. The most recent frontier concerns generative AI—large language models capable of drafting, summarizing, and personalizing sustainability reports. Generative AI impacts sustainability reporting across the value chain: automated drafting of report sections from structured data inputs; summarization of lengthy disclosures for specific stakeholder groups; and personalization of disclosure delivery to match diverse audience needs[3]. For corporate reporting functions facing CSRD compliance deadlines, generative AI reduces the labor intensity of report preparation by automating narrative drafting from quantitative metrics. However, significant risks accompany these capabilities: AI-generated text may be misused to produce misleading narratives, amplifying greenwashing risks, and reliance on incomplete or biased training data may lead to inaccurate predictions, compromising credibility[3].

4 Emerging Tensions and Implications

The cumulative effect of AI applications is a fundamental reconfiguration of ESG disclosure from a narrative, self-reported practice toward a data-driven, verifiable, and increasingly mandatory transparency system. AI technologies progressively transform each stage of the ESG disclosure value chain: NLP-powered textual analysis outperforms manual content analysis in scalability and reproducibility; interpretable ML frameworks enhance predictive accuracy and transparency; and Audit 4.0 tools enable assurance providers to independently triangulate company-reported environmental data against exogenous sources[4]. Despite this transformative potential, effective AI deployment is contingent on resolving persistent tensions around data quality and standardization. The simultaneous operation of diverse reporting standards—GRI, SASB, TCFD, ESRS, and ISSB—introduces variability that undermines the stability and comparability of ML-driven ESG scoring models. ESG data are noisy, incomplete, inconsistent across providers, and subject to measurement errors that are not randomly distributed[19]. The effectiveness of ML models rests directly on the quality and comparability of underlying disclosure data. Progressive standardization under ISSB and CSRD frameworks is likely to enhance the reliability of AI-based analytics by providing more uniform input data, but full convergence remains distant. A second tension concerns greenwashing. While AI tools can detect misleading claims, they also create new vectors for greenwashing. Generative AI can produce polished sustainability narratives lacking substantive backing, and AI-generated reports may reproduce systematic biases, prioritizing metrics susceptible to manipulation[3]. Algorithmic greenwashing—using AI to automate misleading disclosures—represents a governance challenge that regulatory frameworks have yet to adequately address. A third tension arises from persistent divergence of ESG ratings across providers. The average pairwise correlation between ESG ratings from six prominent agencies is only 0.54, compared with 0.92 for credit ratings[20]. This divergence decomposes into three sources: scope (differences in measured attributes), measurement (differences in how attributes are measured), and weight (differences in aggregation). Measurement divergence contributes 56% of total divergence, followed by scope at 38% and weight at 6%[20]. Greater ESG disclosure actually increases rating divergence, as more information provides agencies with greater latitude to apply distinct methodologies. AI may either exacerbate or ameliorate this divergence. If different AI models are trained on different data sources or objective functions, they may amplify measurement divergence. Conversely, if AI tools are deployed on standardized, machine-readable disclosure data mandated by CSRD digital tagging requirements, they may reduce scope divergence by ensuring all rating models process the same underlying information. Finally, the transformational potential of AI is mediated by organizational capability and governance. Organizational AI adoption significantly enhances the comprehensiveness and credibility of sustainability disclosures, but this relationship is mediated by dynamic capability development and sustainability committee heterogeneity. Organizational AI capability, rather than mere tool availability, emerges as the key determinant of disclosure quality improvement. This aligns with dynamic capability theory and the resource-based view: firms that invest in AI capabilities as integrated organizational assets achieve systematic improvements in disclosure quality, whereas firms deploying AI merely as point solutions generate superficial compliance behaviors. For institutional investors, the scalability of NLP and ML tools enables processing vast quantities of sustainability disclosures across large portfolios, identifying risks invisible to traditional approaches. However, investors must navigate AI-generated greenwashing, rating divergence, and model bias. They should prioritize AI tools trained on standardized, regulatory-aligned data sources. For regulators, the CSRD requirement for digital, machine-readable disclosure provides infrastructure for AI-powered regulatory technology, enabling automated compliance screening. However, oversight must address algorithmic bias, model opacity, and the environmental footprint of AI training. The IAASB's ISSA 5000 standard represents an important step toward institutionalizing technology-enabled assurance, but further guidance is needed on the governance of AI models used in disclosure preparation and verification. For corporate practitioners, AI integration into ESG reporting is becoming a strategic imperative. Firms must invest in both ESG data collection infrastructure and the organizational capabilities required to deploy AI effectively. Compliance strategies focused exclusively on software procurement are likely to fail. Instead, firms should develop integrated technological ecosystems in which AI operates as the analytical layer of a broader data governance architecture.

5 Conclusion and Future Research

The convergence of escalating ESG disclosure mandates and AI advances is transforming sustainability reporting from self-reported narratives into a data-driven, verifiable system. This paper analyzed this shift across NLP-based text mining, ML-driven ESG scoring, AI-enabled assurance, and the regulatory and organizational conditions shaping adoption. We find that realizing AI's transformative potential hinges on overcoming tensions around data quality, standardization, greenwashing, and ESG rating divergence, while organizational AI capability—not mere technology acquisition—drives improvements in disclosure quality. Critical research priorities include developing multilingual NLP models to reduce English-language bias, leveraging natural experiments (e.g., staggered CSRD implementation) to establish causal effects, advancing AI governance standards in assurance through field experiments, adapting tools to underrepresented emerging markets, and interrogating the normative implications of delegating credibility judgments to algorithms.

Funding

This work was supported by the Research Fund of Chengdu Jincheng College "Digital Economy and Real Economy Integration Research" (NO: 2025JCKY(B)0016)

References

[1] European Commission. (2022). Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. Official Journal of the European Union. https://eur-lex.europa.eu/eli/dir/2022/2464.

[2] IFRS Foundation. (2023). IFRS S1 General Requirements for Sustainability-related Disclosures and IFRS S2 Climate-related Disclosures. https://www.ifrs.org/issued-standards/ifrs-sustainability-standards-navigator/.

[3] De Villiers C., Dimes R., & Molinari M. (2024). How will AI text generation and processing impact sustainability reporting? Critical analysis, a conceptual framework and avenues for future research. Sustainability Accounting, Management and Policy Journal, 15(1):96–118.

[4] Olivares P. A. V., & Artorga de la Cruz, B. J. (2026). AI and data analytics in sustainable financial reporting and ESG disclosure: A systematic literature review. Sustainability, 18(11):5393.

[5] Higgins C., Tang S., & Stubbs W. (2020). On managing hypocrisy: The transparency of sustainability reports. Journal of Business Research, 114, 395–407.

[6] Fiechter P., Hitz J.-M., & Lehmann N. (2022). Real effects of a widespread CSR reporting mandate: Evidence from the European Union's CSR directive. Journal of Accounting Research, 60(4):1499–1549.

[7] Smeuninx N., De Clerck B., & Aerts W. (2020). Measuring the readability of sustainability reports: A corpus-based analysis through standard formulae and NLP. International Journal of Business Communication, 57(1):52–85.

[8] Chabot M. (2025). Textual and AI-based analysis of climate disclosures: Evidence from the European energy sector. The British Accounting Review, 57, 101736.

[9] Stammbach D., Webersinke N., Bingler J., Kraus M., & Leippold M. (2023). Environmental claim detection. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 1051–1066). Association for Computational Linguistics.

[10] Moodaley W., & Telukdarie A. (2023). Greenwashing, sustainability reporting, and artificial intelligence: A systematic literature review. Sustainability, 15(2):1481.

[11] Sahu A. K., Debata B., & Dash S. R. (2024). Manager sentiment, policy uncertainty, ESG disclosure and firm performance: A large language model in corporate landscape. International Journal of Accounting Information Management, 32(5):858–882.

[12] D'Amato V., D'Ecclesia R., & Levantesi S. (2024). Firms' profitability and ESG score: A machine learning approach. Applied Stochastic Models in Business and Industry, 40(2):243–261.

[13] Yang J., Niu Y., Shi W., & Zhu K. (2025). Predicting ESG disclosure quality through board secretaries' characteristics: A machine learning approach. Research in International Business and Finance, 76, 102865.

[14] Subramaniam R. K., Samuel S. D., Seera M., & Alam N. (2024). Utilising machine learning for corporate social responsibility (CSR) and environmental, social, and governance (ESG) evaluation: Transitioning from committees to climate. Sustainable Futures, 8, 100329.

[15] Crocco E., Broccardo L., Alofaysan H., & Agarwal R. (2024). Sustainability reporting in carbon-intensive industries: Insights from a cross-sector machine learning approach. Business Strategy and the Environment, 33(7):7201–7215.

[16] Dai J., & Vasarhelyi M. A. (2016). Imagineering audit 4.0. Journal of Emerging Technologies in Accounting, 13(1):1–15.

[17] Gu Y., Dai J., & Vasarhelyi M. A. (2023). Audit 4.0-based ESG assurance: An example of using satellite images on GHG emissions. International Journal of Accounting Information Systems, 50, 100625.

[18] Li N., Kim M., Dai J., & Vasarhelyi M. A. (2024). Using artificial intelligence in ESG assurance. Journal of Emerging Technologies in Accounting, 21(2):83–99.

[19] Kotsantonis S., & Serafeim G. (2019). Four things no one will tell you about ESG data. Journal of Applied Corporate Finance, 31(2):50–58.

[20] Berg F., Kölbel J. F., & Rigobon R. (2022). Aggregate confusion: The divergence of ESG ratings. Review of Finance, 26(6):1315–1344.