

Article

AI-Enabled ESG Compliance Audit for Stakeholders

Eid M. Alotaibi ^{1,*}  and Abdulaziz M. Alwathnani ² ¹ Department of Accounting and Information Systems, American University of Sharjah, Sharjah 26666, United Arab Emirates² Department of Accounting, Al Yamamah University, Riyadh 11512, Saudi Arabia; a_alwathnani@yu.edu.sa

* Correspondence: ealotaibi@aus.edu

Abstract

Environmental, social, and governance (ESG) disclosures face credibility risks due to Scope 2 Greenhouse Gas (GHG) reports lacking standardized compliance checks, raising concerns about their reliability. This study therefore develops and evaluates an AI-enabled artefact for ESG compliance auditing. This artefact applies natural language processing (NLP) to extract reported values, implements rule-based checks grounded in the GHG Protocol, and produces transparent output. A design science research (DSR) approach guided the design, demonstration, and evaluation of the artefact, which was applied to sustainability reports from five technology companies. The results revealed that it replicates auditor judgments and reduces workload by over ninety percent in the sample. These findings serve as a proof-of-concept for automation in ESG compliance auditing. The theoretical contributions include extending the literature on AI in ESG auditing by reframing its role from producing interpretive scores to enabling transparent compliance verification. This study also demonstrates how DSR can help produce artefacts that embed rule-based logic into ESG assurance with rigor and practical relevance. The practical contributions include highlighting how a lightweight tool can enable auditors, regulators, boards, and investors to screen disclosures and benchmark credibility without sacrificing professional judgment.

Keywords: artificial intelligence; natural language processing; scope 2 emissions; compliance audit; ESG audit; design science research



Academic Editors: Ahreum Hong
and Yannan Li

Received: 24 August 2025

Revised: 14 October 2025

Accepted: 21 October 2025

Published: 25 October 2025

Citation: Alotaibi, E.M.; Alwathnani, A.M. AI-Enabled ESG Compliance Audit for Stakeholders. *Sustainability* **2025**, *17*, 9513. <https://doi.org/10.3390/su17219513>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Environmental, social, and governance (ESG) disclosures have become central to evaluating corporate accountability and sustainability performance. Regulators and investors now demand more comprehensive information about Greenhouse Gas (GHG) emissions, supply chain practices, and societal impacts. Despite some progress, credibility concerns persist due to many ESG reports lacking standardized compliance checks. Research has shown that firms often report inconsistent values, particularly for Scope 2 GHG disclosures, where location-based and market-based values can differ significantly [1–3].

The reporting of carbon emissions under the GHG Protocol is typically divided into three categories: Scope 1, Scope 2, and Scope 3. Scope 1 encompasses the direct emissions from company-owned operations, while Scope 2 covers the indirect emissions from purchased energy. Finally, Scope 3 includes any other indirect emissions across the value chain. This study focuses on Scope 2 emissions because they account for a large share of the typical corporate footprint and are directly linked to electricity consumption. They are also subject to methodological differences in location-based and market-based reporting [3]. By focusing on this category, this study addresses one of the most critical yet inconsistent areas

of ESG disclosure. Without systematic assurance, stakeholders (stakeholders in the context of ESG reporting include parties who rely on disclosed information for decision-making, such as investors, regulators, boards of directors, auditors, vendors, creditors, customers, and civil society groups) face delayed decisions and uncertainty about whether reported information reflects actual performance.

Auditors play a critical role in allaying such concerns, but the profession faces structural challenges. Indeed, expanding the requirements for sustainability disclosure in turn increases the scope of the necessary assurance engagements, while a global shortage of qualified auditors constrains capacity [4]. Such pressures are exacerbated by short reporting timelines and complex ESG metrics. Traditional manual reviews involve auditors locating dispersed disclosure sources, extracting the relevant data, and testing for compliance under tight deadlines. This leads to the risk of inconsistencies and misclassifications arising that could reduce stakeholders' confidence in ESG reporting.

Artificial intelligence (AI) offers a potential solution by automating repetitive aspects of the assurance process while safeguarding the need for auditor judgment. Prior studies have shown that digital tools can increase transparency and reduce credibility risks in reporting contexts [5,6]. Empirical applications of AI to ESG compliance auditing remain limited, however. Indeed, existing work has mostly focused on blockchain for transparency or broad AI applications in auditing rather than operationalizing compliance rules into a functioning artefact for ESG assurance. This gap persists in both research and practice, underscoring the need for transparent artefacts that enable consistent compliance testing.

This study addresses this gap by developing and evaluating an AI-enabled artefact for ESG compliance auditing. This artefact extracts Scope 2 GHG disclosures, applies standardized rules based on the GHG Protocol, and produces output that can be interpreted by auditors and stakeholders. The design science research (DSR) paradigm was adopted, with the resulting artefact being applied to sustainability reports from technology firms and evaluated against manual reviews.

The objective of this study was to establish how AI can enhance ESG compliance auditing by improving efficiency, consistency, and transparency. To guide this investigation, this study followed two research questions:

RQ1: *How can an AI-enabled artefact operationalize compliance rules to test ESG disclosures consistently?*

RQ2: *How do the results of such an artefact compare with manual compliance reviews in terms of accuracy and efficiency?*

This study is structured as follows: Section 2 reviews the relevant literature on ESG reporting, assurance, and AI in auditing. Section 3 presents the motivation driving the study and its expected contribution. Section 4 describes the research methodology, including the design, demonstration, and evaluation of the artefact. The study's findings and their implications for practice, governance, and research are then discussed in Section 5. Finally, Section 6 concludes this paper and outlines some potential avenues for future research.

2. Literature Review

ESG reporting has shifted from being a voluntary disclosure to becoming a mandatory component of corporate accountability. Global frameworks now shape the scope and comparability of sustainability information, and while assurance practices have expanded, they remain fragmented. Research in auditing and information systems highlights AI's potential to improve verification, while blockchain and other digital trust technologies enhance transparency and traceability. Investors, regulators, and boards rely on ESG data to assess risk and evaluate governance, yet current systems continue to produce inconsistent

reports with opaque assurance. This section reviews relevant key areas for auditing ESG compliance and highlights the research gap that motivated this study.

2.1. ESG Reporting and Compliance Frameworks

ESG disclosure is governed by multiple global frameworks that define scopes, metrics, and integration with financial reporting. The Global Reporting Initiative (GRI) remains the most widely adopted framework, with it offering topic-specific indicators to help organizations measure and convey their sustainability performance [7]. The Sustainability Accounting Standards Board (SASB), meanwhile, complements the broader frameworks by providing industry-specific metrics that emphasize financial materiality. Analyzing this framework offers significant insight into the robustness of sustainability reporting approaches and how they can be tailored to the needs of diverse stakeholders. Investors, civil society organizations, and regulators, who actively monitor and critique corporate disclosures, can gain a deeper understanding of a phenomenon when reporting is aligned with SASB standards, which in turn enhances their ability to evaluate risk, accountability, and long-term performance [7].

The International Sustainability Standards Board (ISSB) introduced IFRS S1 and S2 to establish a global baseline for sustainability disclosure based on integrating climate and ESG data into general-purpose financial reports [8]. These frameworks together aim to enhance comparability, transparency, and audit readiness. The introduction of the European Union's Corporate Sustainability Reporting Directive (CSRD) represented a regulatory turning point by mandating ESG disclosures for large firms. More specifically, the CSRD requires application of the European Sustainability Reporting Standards (ESRS) and introduces explicit assurance obligations [9]. This regulatory step shifted disclosure from a voluntary to an enforceable compliance practice, so it now requires firms to reconcile narratives with verifiable evidence. Investors and regulators therefore have more authority to demand consistency and accuracy in reported data.

GHG reporting is central to these frameworks, particularly under the GHG Protocol. This protocol requires Scope 2 emissions to be disclosed using both location-based and market-based methods [10]. More specifically, location-based values reflect average grid emission factors, while market-based values reflect contractual arrangements, such as power purchase agreements and renewable energy certificates. The intention is to give stakeholders a full picture of both the physical emissions and the contractual positioning. Nevertheless, research has shown that the difference between the two values can be very substantial, sometimes by over 90 percent, with market-based reporting often understating carbon intensity when certificates are used instead of actual grid decarbonization [1,3]. More recent studies have confirmed that post-2020 Scope 2 disclosures still exhibit wide variations across regions and industries, underscoring the absence of practical benchmarks for assessing plausibility [11,12].

Despite their global adoption, the current frameworks remain limited in addressing credibility concerns. While they define disclosure requirements, they stop short of prescribing thresholds or compliance tests for validating the reported values. Consequently, firms can present very low market-based emission values without clearly justifying them, while stakeholders lack standardized methods for evaluating whether the reported values are plausible. Such weaknesses expose sustainability reporting to the risk of inconsistency and selective presentation. This shortcoming motivates the development of automated tools that operationalize compliance rules in a transparent and replicable manner.

2.2. Assurance and Audit of ESG Disclosures

External assurance has gained popularity as a means for strengthening the credibility of ESG reports. While early studies have shown that assurance can add legitimacy to sustainability disclosures, they also highlighted wide variations in provider types, applied standards, and assurance scope [13]. More recent work has also found that the type of assurer influences quality: For example, accounting firms tend to focus on selected metrics using structured methodologies, while non-accounting providers often cover broader topics with less rigor [13]. This variation results in assurance statements that differ in depth and transparency, limiting any value they have for stakeholders [14].

Nevertheless, assurance contributes to governance by improving a board's oversight of sustainability performance [15], yet the effectiveness of this oversight is critically dependent on the quality of the underlying data. For example, auditors face challenges when ESG information originates from fragmented systems and inconsistent data sources, potentially compromising the reliability of the assurance [16]. Scholars argue that without regulatory enforcement and harmonized criteria, assurance is unlikely to deliver consistent value to investors [17]. Market incentives also shape choices for assurance, and in some cases, these can result in symbolic or superficial audits rather than any substantive testing of disclosures [18].

The regulatory environment is beginning to address these shortcomings, however. The CSRD requires assurance for sustainability reporting in Europe, while the ISSB's standards anticipate a global alignment that will grow the demand for audit-like services [19]. The current practices remain heterogeneous in nature, though. Providers apply differing levels of assurance ranging from limited assurance under ISAE 3000 [20] to bespoke criteria or local standards [21]. Such inconsistencies all but exclude any comparisons across firms and markets.

Another persistent challenge lies in the methodology itself: Traditional assurance relies on manual reviews, interviews, and document testing, but such methods are resource-intensive. What is more, constrained reporting timelines limit the ability to test numerical plausibility or reconcile disclosures with external benchmarks. Sample-based approaches can increase efficiency but at the risk of overlooking systematic errors in emissions data [22]. In the absence of any automated compliance checks, assurance remains dependent on human judgment, thus raising concerns about scalability as disclosure requirements expand under the CSRD, ISSB, and other frameworks.

As a whole, the literature reinforces that assurance adds value for legitimacy and governance, but it falls short of providing consistent, scalable, and transparent compliance testing. There is therefore an opportunity for technological solutions, particularly AI, to automate routine verification tasks, thus reducing costs while complementing professional judgment rather than replacing it.

2.3. AI in Auditing and Compliance

The limitations of manual assurance, as outlined above, have led to a growing interest in applying AI to support audit and compliance processes. Manual assurance relies heavily on sampling and professional judgment under strict time constraints, which together limit coverage and consistency. AI offers a potential pathway for expanding the reach of audits by processing large volumes of unstructured information and generating transparent output that can be applied consistently across firms [23].

Applications of AI in financial auditing have already demonstrated the value of this approach. For example, anomaly-detection models can identify unusual journal entries that signal elevated risk [24]. Natural language processing (NLP) systems, meanwhile, can extract contractual obligations from lengthy documents more accurately than a traditional

manual review [25]. Predictive analytics add further coverage by highlighting transactions or accounts that warrant closer scrutiny [24]. Such applications improve efficiency by directing auditors' attention to areas where their professional judgment will be most critical.

Within the sustainability domain, AI has been mostly applied to disclosure classification, ESG ratings construction, and sentiment evaluation. For instance, NLP has been used to assess the tone of sustainability reports and categorize disclosures into standard themes [26]. Machine learning models have also been applied to create composite ESG scores that investors can use as proxies for firm performance [27]. More recent studies have also evaluated AI's role in assessing disclosure credibility, including work on demonstrating how AI can improve sustainability reporting in sector-specific contexts [22,27]. While no doubt valuable in their own right, these approaches do not evaluate the numerical plausibility of reported emissions. Indeed, they provide interpretive scores rather than conducting rule-based compliance checks.

This gap becomes particularly evident for Scope 2 emissions, where the GHG Protocol mandates the disclosure of both location-based and market-based values. The considerable divergence between the two methods raises concerns for stakeholders, yet there is no automated mechanism to test whether the disclosed values fall within a reasonable threshold [3]. The current AI tools are not designed to benchmark market-based numbers against residual mix factors or to flag cases where contractual positioning produces implausible outcomes. Moreover, the reliance on opaque or "black-box" models, where AI systems produce results without revealing the underlying rules or decision logic, introduces risks like extraction errors, a lack of transparency, and a potential overreliance on algorithmic outputs. This further underscores the need for systems that embed clear, rule-based logic into their design [27].

The absence of transparent, rule-based artefacts for ESG assurance represents a critical gap. Indeed, black-box ESG ratings are often criticized because the criteria being used to generate scores are not disclosed, thus limiting their usefulness for compliance or regulatory purposes [28]. A transparent, rule-based AI system that applies defined checks for Scope 2 disclosures would therefore represent a significant advance. By moving beyond interpretive scoring toward compliance verification, such an artefact could directly support stakeholders in evaluating the credibility of emission reports.

2.4. Digital Trust Technologies and Transparency

Digital trust technologies represent another stream of work that seeks to address weaknesses in ESG reporting. Alotaibi, Issa and Codesso [5] designed a blockchain-based model for government records using a DSR approach, with them showing how this can strengthen transparency and auditability. Alotaibi, Khallaf, Abdallah, Zoubi and Alnesafi [6] extended this idea to carbon accountability in supply chains and demonstrated how distributed ledgers can serve as verifiable records of emissions claims. Other studies have yielded similar findings, such as that of Saberi et al. [29], which explored blockchain applications in sustainable supply chains. Casino et al. [30], meanwhile, analyzed how distributed ledgers could improve auditing and compliance processes. The above contributions underscore the value of blockchain for creating tamper-resistant records.

Despite such advances, blockchain solutions primarily aim to address traceability and record integrity. As such, they confirm that transactions or reported claims have not been altered, but they do not ascertain whether the disclosed values are plausible or consistent with external benchmarks. In other words, blockchain secures the history of information, but it does not assess the credibility of the original numbers. This limitation means that despite the stronger transparency, stakeholders still need automated mechanisms to test whether emission disclosures align with expected physical performance.

The gap left by blockchain-based systems therefore highlights the need for complementary AI-enabled artefacts. While blockchain improves data integrity, AI can support assurance by applying compliance rules directly to the reported figures. The present study positions its artefact within this nexus by leveraging the transparency benefits of digital trust systems while addressing the as-yet-unmet need for the automated compliance testing of ESG disclosures [6]. This study extends prior work by positioning AI artefacts as being complementary to blockchain for addressing ESG reporting credibility through automated compliance audits.

2.5. Stakeholder Use of ESG Information

Investors rely on ESG disclosures to evaluate value and exposure to risk. Firms with material ESG reporting show stronger financial performance when their disclosures align with recognized standards, thus reinforcing the economic significance of credible sustainability information [31]. Comparability remains a persistent challenge, however, and investors can be skeptical about the reported progress when disclosures differ in scope or measurement approaches [28]. For investment decision-making, the lack of transparent compliance checks increases the risk of misallocating capital to firms that merely appear sustainable through selective reporting.

Regulators are advancing standardization to protect market integrity, however. Initiatives like the ISSB's global baseline and the European CSRD require firms to integrate sustainability reporting into their financial statements under strict assurance obligations [19]. Such measures improve accountability, but they also expose regulators to the need to monitor thousands of reports across diverse industries. Automated tools that screen disclosures for internal consistency and plausibility could help regulators to focus their limited resources on outliers rather than casting a wide net.

Corporate boards also see ESG reporting as being integral to governance. Reliable disclosures provide a basis for overseeing sustainability strategies and allowing boards to get ahead of regulatory scrutiny. Nevertheless, with assurance engagements being resource-intensive and limited in scope, boards lack a timely mechanism to verify whether disclosures are likely to withstand external scrutiny [15]. NGOs and civil society groups also depend on ESG disclosures to detect greenwashing and hold firms accountable, and they also face challenges in verifying the reported data at scale.

The needs of stakeholders underscore a gap in current assurance practices. Frameworks and assurance standards emphasize disclosure while providing limited tools for testing compliance with benchmarks, such as residual mix factors. This results in stakeholders depending on selective assurance statements or complex reports that may obscure anomalies. The DSR provides a pathway for addressing this gap by developing an artefact to operationalize the rules transparently. Indeed, prior applications of DSR have shown how digital artefacts can improve auditability and accountability in sustainability contexts [5,6]. Applying DSR to ESG compliance auditing will enable the creation of an AI system for directly testing the credibility of reported numbers and delivering outputs that are transparent, replicable, and aligned with the needs of multiple stakeholders.

3. Motivation and Contribution

3.1. Research Gap

The motivation for this study arose from the combination of mandatory sustainability reporting, the growing demand for credible assurance, and the absence of transparent, automated tools for compliance testing. Standards like GRI, SASB, ISSB, and CSRD require firms to disclose detailed ESG information without providing mechanisms to verify the plausibility or internal consistency of the reported values. The phenomenon of external

assurance has grown, but it remains costly, inconsistent, and dependent on the scope of individual providers. Regulators face difficulties in enforcing compliance across thousands of reports, while investors rely on disclosures that often contain discrepancies between contractual claims and physical performance. This situation motivates the design of an artefact that automates compliance process in a way that is simple, transparent, and aligned with established standards.

The problem is illustrated most vividly in the case of Scope 2 GHG reporting. The GHG Protocol requires disclosure of both location-based and market-based emissions, with the former reflecting average grid intensity, while the latter incorporates contracts and renewable energy certificates purchased by organizations. This dual view theoretically helps stakeholders by distinguishing physical emissions from contractual positioning, but in practice, large gaps emerge between the two methods. Research has shown that some firms report market-based emissions far below their location-based equivalent, sometimes by more than 90 percent, creating an impression of rapid decarbonization that is not reflected in the actual grid mix [1,3]. Such discrepancies hinder a meaningful interpretation, because investors cannot easily determine whether the progress is real or the result of some accounting instrument. Auditors also face the challenge of testing whether contracts fulfil their claims, often without access to independent verification data. Regulators also rely on voluntary disclosure notes that may not adequately explain the scale of the gap. This recurring issue provides a clear use case for automated compliance checks to support assurance.

A second dimension of the problem lies in plausibility: Even when firms disclose both market-based and location-based emissions, there is little guidance about what constitutes a reasonable difference. For example, the GHG Protocol mandates disclosure without prescribing thresholds for evaluation. This allows firms to present very low market-based values thanks to contractual certificates, even if the values diverge substantially from the residual mix reported at the regional or national level. Without a reference point, stakeholders lack a way to test whether the reported numbers are credible. A compliance auditing artefact could compare the reported market-based values with residual mixed expectations and flag any excessive deviations. Such a tool would not fully replace a manual assurance, but it could serve as a rapid, transparent initial screening for directing expert attention to high-risk disclosures.

This study's motivation also connects directly with the theme of business sustainability and efficiency. ESG compliance has shifted from a voluntary reporting practice to a mandated requirement that underpins corporate accountability [32]. Auditing also plays a central role in maintaining public trust, yet traditional assurance methods rely on materiality judgments and sampling approaches that are constrained by regulatory deadlines [33]. Such constraints limit the depth and consistency of compliance testing, raising concerns about whether the reported information accurately reflects the underlying performance. Advanced technologies, particularly AI, offer ways to overcome these challenges by automating routine checks, reducing assurance costs, and expanding the scope of a review within the same reporting window [34]. By applying automation to the first stage of compliance review, the artefact developed in this study reduces the manual burden, increases the likelihood of detecting anomalies, and strengthens the credibility of the reported outcomes. Applying such advanced technology to auditing will support sustainable business practices by ensuring that disclosures reflect actual performance rather than accounting instruments. It will also enable precious auditing resources to be directed toward high-risk areas where their professional judgment is most needed [23].

3.2. Contributions

From a theoretical perspective, the study advances research on AI in auditing by demonstrating how rule-based automation can be applied to ESG compliance testing. Prior work has focused on AI applications for ratings construction, sentiment evaluation, or disclosure classification, which provide interpretive scores rather than transparent verification of reported values [22,27,35]. By contrast, this study develops a functioning artefact that operationalizes the GHG Protocol to test the plausibility of Scope 2 disclosures. This extends the literature by demonstrating how DSR can embed compliance rules into functioning artefacts and reframe the role of AI in ESG auditing from interpretive scores to compliance verification.

From a practical perspective, the artefact provides a lightweight compliance tool that stakeholders can apply in different ways. Auditors can reduce the burden of manual extraction and direct their attention to high-risk areas. Regulators can also use the tool to focus enforcement on firms that report implausible discrepancies. Boards can test the credibility of their own disclosures before external assurance, while investors can benchmark reported values against transparent rules to identify credibility risks. These applications show how automation can strengthen the efficiency, comparability, and reliability of ESG assurance while leaving interpretation to professional judgment.

4. Methodology

This study applied the DSR paradigm to create and evaluate an artefact for ESG compliance auditing. DSR is appropriate when research seeks both a practical solution and a theoretical contribution [36]. This method proceeds through the phases of problem identification, objective setting, artefact design, demonstration, and evaluation. Each phase helps refine the resulting artefact and generates insights for academic and professional audiences.

4.1. Problem Identification

Section 3 outlined the desire to address weaknesses in Scope 2 reporting. Such disclosures often feature large differences between the location-based and market-based values, but stakeholders lack any consistent method to ascertain whether those differences are credible. Location-based values reflect the average emissions intensity of the electricity grid where consumption occurs, independent of contractual purchases, while market-based values are calculated based on supplier contracts and renewable energy certificates acquired by firms. The GHG Protocol requires firms to report both values to provide a complete view of physical electricity use and contractual positioning [3], but it does not specify compliance thresholds or checks. Consequently, the reported market-based values may differ drastically from the residual mix factors without explaining why. This diminished trust in sustainability reporting ultimately leaves investors, boards, and regulators uncertain about whether they can interpret the results with confidence.

Furthermore, the absence of automated testing tools compounds the issue. External assurance remains costly and heterogeneous, and manual reviews of reports are slow and subjective. In addition, large volumes of sustainability disclosures make it impractical for regulators to enforce consistency at scale. Investors also lack mechanisms for screening portfolios for potential outliers in a systematic way. All this leads to the problem that this study addresses: ESG reporting lacks transparent, scalable compliance checks that can ascertain numerical plausibility and internal consistency.

4.2. Objectives of the Artefact

The artefact was designed to deliver transparent compliance checks for Scope 2 disclosures. Its primary objective is therefore to extract reported values from sustainability

reports in a structured and reliable manner. Scope 2 data are often presented in tables or embedded in narrative text, and any extraction must capture both location-based and market-based values accurately. Once extracted, the artefact applies compliance rules that align directly with the GHG Protocol. The first rule measures the relative difference between the two Scope 2 calculation methods, while the second rule compares reported market-based values with the residual mix expectations. Both rules are grounded in prior research that has highlighted the risk of unexplained gaps and implausible values [1–3].

A secondary objective is to present output in a form that stakeholders can interpret without technical training. The artefact therefore produces compliance profiles for each firm. Flags are clearly indicated, with each result being supported by the formula that generated it. The artefact is also modular, so new compliance rules can be added to cover other ESG disclosures, such as Scope 3 emissions, water usage, or diversity metrics. Efficiency is another secondary goal, because the system must reduce the time required for a compliance review while maintaining accuracy. In meeting these objectives, the artefact will serve as a proof of concept for automating ESG compliance auditing while also providing a platform for future expansion.

4.3. Artefact Framework

The framework follows a sequential process from data collection to output generation while also incorporating bi-directional flows that reflect verification and error-checking routines. Figure 1 illustrates the main components of the system for transforming sustainability reports into compliance profiles.

The process begins by collecting relevant data. For this study, corporate sustainability reports in PDF or HTML format were used as the primary input. These reports often contain Scope 2 disclosures in tabular or narrative form, so the artefact applies NLP techniques to extract the relevant text. More specifically, text mining is used to search for keywords like “Scope 2”, “location-based”, “market-based”, and “electricity consumption”, with these patterns being linked to numerical expressions. Cleaning routines standardize the formatting, remove non-textual elements, and harmonize units to ensure comparability across reports with different original layouts.

The extraction stage serves as the central hub for linking raw reports with structured data. This is depicted in Figure 1 with bi-directional arrows because extraction is not a one-way process. When the initial parsing detects incomplete or ambiguous results, the artefact loops back to reprocess the source report or cross-check the extracted values against the structured dataset. This ensures that missing or misaligned disclosures are identified transparently rather than silently omitted.

The NLP pipeline combines text mining with contextual analysis to ensure the accurate identification of Scope 2 data. After parsing the reports into machine-readable text, the artefact applies rule-based keyword detection for terms like “Scope 2”, “location-based”, “market-based”, and “electricity consumption.” These keywords are cross-referenced with numerical expressions in the surrounding context to isolate candidate values. To avoid misclassifying them, named entity recognition and dependency parsing confirm that the extracted numbers correspond to GHG disclosures rather than unrelated metrics, such as financial figures. This layered approach of combining keyword detection, contextual parsing, and entity recognition reduces the risk of false positives and increases confidence in the extracted LB, MB, and electricity consumption values [37].

Once extraction is complete, the framework captures the required variables, namely location-based (LB) emissions, market-based (MB) emissions, and electricity consumption where disclosed. These values are necessary for applying the compliance rules. Electricity consumption figures are further combined with residual mix factors to generate an “ex-

pected” MB benchmark. Residual mix factors are published annually by the Association of Issuing Bodies (AIB) for Europe and equivalent bodies for North America. These factors reflect the emissions intensity of electricity that is not covered by renewable contracts, making them an appropriate baseline for plausibility testing.

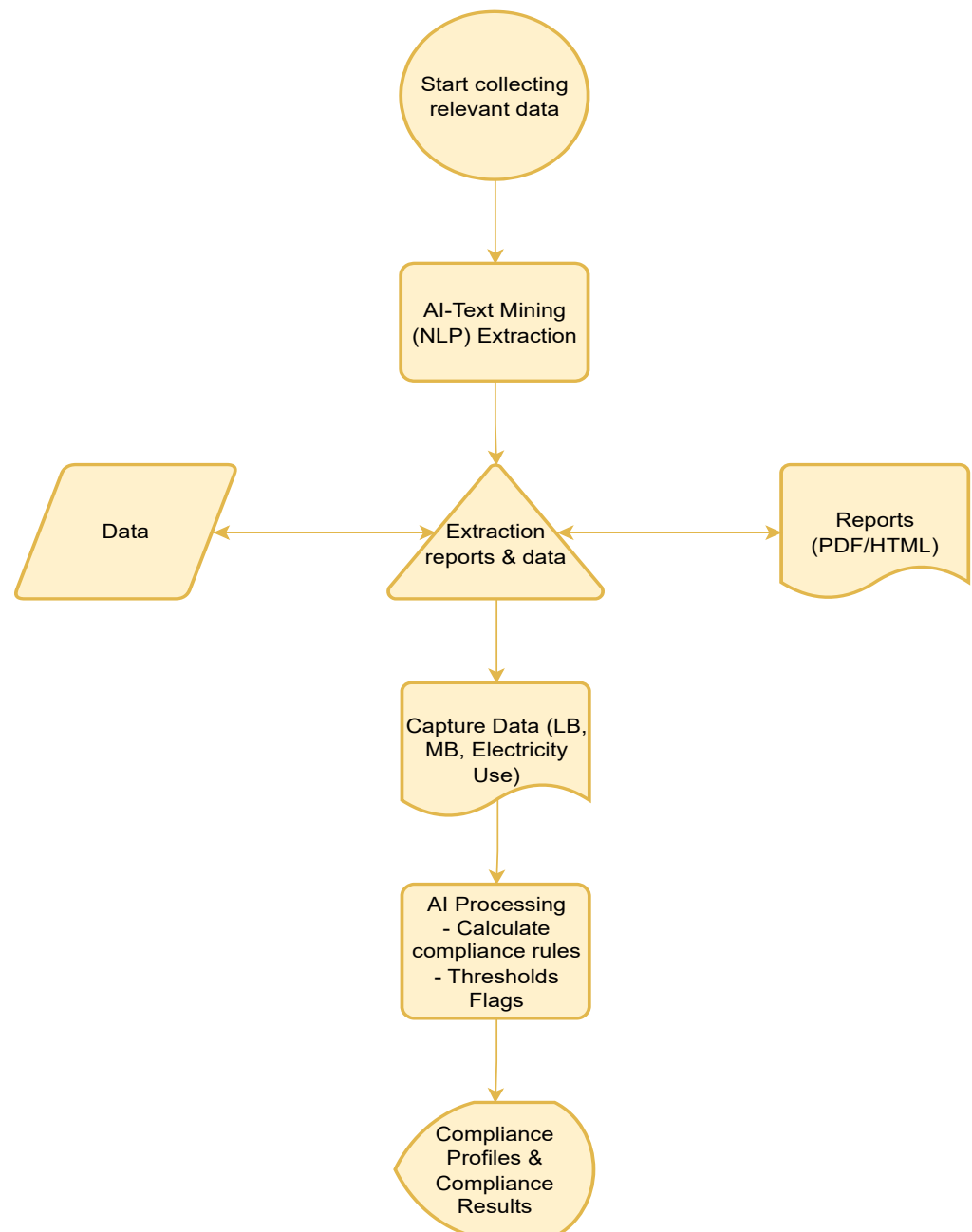


Figure 1. The AI-enabled artefact framework for ESG compliance auditing.

The next stage is the AI-based processing, where the rule engine calculates compliance metrics and applies thresholds. The gap rule measures the relative difference between the LB and MB emissions, while the plausibility rule compares the MB values against residual mix expectations. The artefact then flags any firms that breach either threshold, so the results will be transparent and easy to interpret.

Finally, the system produces compliance profiles and outputs. These include the extracted values, calculated metrics, and rule-based flags. The reports are stored in structured tabular form, making them easily reproducible and suitable for audit documentation. This ensures that stakeholders—including investors, boards, regulators, and

assurance providers—can access transparent, reproducible results rather than opaque algorithmic judgments.

These stages illustrate how the artefact transforms complex, heterogeneous sustainability reports into standardized compliance results. The inclusion of a bi-directional flow emphasizes how extraction and validation are iterative processes, and this represents a critical design choice for maintaining data integrity when working with unstructured disclosures.

4.4. Artefact Design

Building directly on the framework illustrated in Figure 1, the artefact's design operationalizes the extraction, rule processing, and output stages into a reproducible system. Each component of the framework was translated into specific routines for capturing Scope 2 data, applying compliance rules, and generating standardized outputs that can be independently verified. The framework begins by applying AI-based text mining to extract relevant Scope 2 data from sustainability disclosures in PDF or HTML documents. More specifically, the documents are converted into structured data for LB values, MB values, and electricity consumption figures. These data are then processed by the artefact's rule engine for two compliance checks, namely the gap rule and the plausibility rule.

The NLP pipeline was designed to handle disclosures in both tabular and narrative formats, because firms may disclose Scope 2 information in either. The process begins by converting PDF or HTML reports into machine-friendly text using document-parsing libraries. Regular expressions and rule-based filters are applied to identify keywords like "Scope 2", "location-based", "market-based", and "electricity consumption", which are then linked to numerical expressions. The NLP model then applies named entity recognition and contextual parsing to enhance accuracy, thus ensuring that the extracted values correspond specifically to GHG disclosures rather than something else. This combination of keyword detection and contextual validation reduces the chance of false matches and improves the accuracy of the extracted LB, MB, and electricity consumption values [37].

The gap rule relates to the relative difference between the LB and MB Scope 2 emissions. LB values are calculated based on the average grid intensity of the consumed electricity, while MB values incorporate contractual arrangements like supplier agreements and renewable energy certificates. A large discrepancy between the LB and MB values could indicate an overreliance on contractual mitigations rather than an actual reduction in physical emissions. The artefact therefore flags instances where the MB value is more than 70 percent below the LB value. This threshold reflects guidance in the GHG Protocol, which requires dual disclosure precisely to highlight such gaps. Differences of this magnitude rarely arise from normal grid conditions and often reflect an aggressive use of contractual instruments over actual decarbonization [1,3,10].

The plausibility rule benchmarks MB values against the expected emissions calculated from residual mix factors, which measure the emissions intensity of the electricity not backed by renewable energy certificates. Expected values are calculated by multiplying electricity consumption by the residual mix factor relevant to the reporting region. The artefact flags cases where this is more than 20 percent, plus or minus, from the expected value. This 20 percent tolerance is consistent with published evidence that has shown that residual mix calculations can typically vary within this margin due to timing mismatches and regional estimation uncertainty. Any reported MB values falling outside this range may raise concerns about credibility because such deviations cannot easily be explained by standard residual mix variability [2,38,39]. In practice, the 20 percent threshold balances the flexibility to account for normal variance with the need to detect any substantial divergence from benchmarks.

The compliance rules were operationalized using transparent formulas. The gap percentage (Equation (1)) is calculated as:

$$Gap\% = \frac{(LB - MB)}{LB} \times 100 \quad (1)$$

The plausibility deviation (Equation (2)) is calculated as:

$$Deviation\% = \frac{(MB - Expected)}{Expected} \times 100 \quad (2)$$

Here, *LB* is the location-based Scope 2 value disclosed in the report, *MB* is the market-based value, and *Expected* is calculated by multiplying the electricity consumption by the residual mix factor of the relevant grid. These equations provide the explicit metrics used by the artefact to generate compliance reports.

Residual mix factors were sourced from publicly available registries to ensure transparency and reproducibility. For Europe, they were taken from those published annually by the AIB, while for North America, equivalent factors were obtained from regional grid authorities and published technical reports. These values represent the emissions intensity of electricity that is not covered by renewable energy contracts, so they serve as a benchmark for plausibility testing. Integrating this data ensured that the expected values are anchored in authoritative data, thus reducing the risk of bias in the compliance check.

Table 1 summarizes the compliance rules and explains the rationale for the chosen thresholds. The 70 percent threshold for the gap rule helps flag cases where MB values appear unrealistically low in relation to LB values. The ± 20 percent tolerance for the plausibility rule, meanwhile, accommodates some residual mix variability while still flagging reported values that diverge substantially from the benchmark value.

Table 1. Compliance rules as applied in the artefact.

Rule	Threshold	Rationale
Gap rule	>70%	Large gaps suggest reliance on certificates without physical reduction.
Plausibility rule	> $\pm 20\%$	Large deviations from residual factors raise credibility concerns.

The artefact was used to extract LB, MB, and electricity consumption values from sustainability reports and then apply compliance rules to generate profiles for each firm. The dataset includes reports from Microsoft, Apple, Google, Amazon, and Meta, all of which are among the world's largest electricity consumers due to their data centers. These firms consistently disclose Scope 2 emissions under both methods together with electricity consumption data, thus enabling a calculation of expected emissions based on published residual mix factors for the relevant region.

4.5. Demonstration

To test the framework shown in Figure 1, the artefact was applied to a dataset of technology firms including Microsoft, Apple, Google, Amazon, and Meta. The artefact applied text mining to capture the Scope 2 disclosures in their corporate reports, with both tabular and narrative depictions being captured consistently and converted into structured data for compliance testing.

This study focused on the technology sector due to the high electricity consumption of data centers around the world, making Scope 2 reports comparable across firms. These firms are among the largest consumers of electricity in the world due to their numerous data

centers, and they consistently publish Scope 2 values for both methods. Moreover, electricity consumption data is also available in their reports, enabling the expected emissions to be calculated based on residual mix factors from relevant regional authorities. By using this dataset, the artefact could be tested in the real-world reporting conditions of the technology sector. For consistency, the extracted reports were drawn from the most recent full-year publications available at the time of analysis: Microsoft’s 2024 Sustainability Report (covering 2023 data), Apple’s 2023 Environmental Progress Report (reflecting 2022 data), Google’s 2023 Environmental Report, Amazon’s 2023 Sustainability Report with associated 2024 verification data, and Meta’s 2023 Sustainability Report (reporting 2022 Scope 2 emissions).

Table 2 presents the compliance audit results for the five firms in the dataset. The artefact applied the gap and plausibility rules to the extracted values and determined whether the thresholds were breached.

Table 2. Compliance audit results for the technology firms.

Firm	Location-Based Scope 2 (tCO ₂ e)	Market-Based Scope 2 (tCO ₂ e)	Gap %	Residual Mix Expected (tCO ₂ e)	Deviation %	Gap Flag	Plausibility Flag
Microsoft	8,200,000	380,000	95%	1,200,000	−68%	Yes	Yes
Apple	5,400,000	2,900,000	46%	3,100,000	−6%	No	No
Google	7,100,000	1,600,000	77%	2,400,000	−33%	Yes	Yes
Amazon	4,800,000	3,900,000	19%	4,200,000	−7%	No	No
Meta	6,500,000	5,700,000	12%	6,100,000	−7%	No	No

Note: All values are reported in metric tons of CO₂ equivalent (tCO₂e).

The compliance audit flagged Microsoft and Google under both rules, so their reported MB values are far below their LB values and deviate substantially from residual mix expectations. Apple’s gap is more moderate, with its MB values remaining within the 20 percent plausibility band, so no flags were raised. Amazon and Meta also reported results within acceptable limits for both rules. These findings confirm the artefact’s ability to identify firms with potential compliance risks while also highlighting reporting inconsistencies that may otherwise be overlooked in a manual review.

Although our evaluation focuses on Scope 2 emissions in the technology sector, the artefact is not sector-dependent. Any sector that routinely discloses location-based values, market-based values, and electricity consumption is a suitable candidate for the extraction and rule engine [12]. For example, aviation, energy, and manufacturing firms publish Scope 2 metrics in formats compatible with the artefact [40]. The modular architecture also enables extension to other ESG disclosures. Water consumption could be benchmarked in relation to regional scarcity factors. Waste intensity could be evaluated in terms of national treatment standards. Diversity metrics could be compared with workforce composition benchmarks. Adjusted rules could perform standardized, transparent checks across multiple ESG areas. The artefact could therefore be applied across sectors without substantial redesign of the core pipeline, though this requires validation. The evaluation, however, was limited to sustainability reports from five companies in the technology sector. This provided consistency in disclosure formats and facilitated application of the artefact, although it also narrowed the scope of the interpretation. As DSR, the evaluation should be regarded as a proof of concept rather than a sector-general result. Broader testing across industries and disclosure categories is required to assess the artefact’s robustness in different contexts.

4.6. Evaluation

The evaluation stage assessed whether the AI-enabled artefact efficiently produces results that are accurate and comparable to a manual compliance review. Accuracy was tested by comparing the artefact's output with benchmark results obtained through manual analysis. Efficiency, meanwhile, was assessed by comparing the time required for each approach.

For accuracy, five participants who were enrolled in a graduate-level advanced auditing course were assigned to manually review one sustainability report each. All participants had at least three years of professional auditing experience prior to enrolling in the course, so they were familiar with practical audit procedures. One participant with four years of experience reviewed Microsoft, a second with three years of experience reviewed Apple, a third with five years of experience reviewed Google, a fourth with three years of experience reviewed Amazon, and a fifth with four years of experience reviewed Meta. Each participant was instructed to extract the Scope 2 disclosures, calculate the gap and plausibility metrics, and evaluate whether the firm met the compliance thresholds.

Table 3 presents the outcomes of the manual reviews alongside the artefact's results. The manual reviews found Apple, Amazon, and Meta to be compliant, while Google was flagged under both rules. Microsoft was initially assessed as being compliant, yet the artefact flagged it under both rules. In follow-up discussions, participants reported difficulty in locating the correct reports and identifying the relevant Scope 2 data, which were often scattered across multiple sections. In Microsoft's case, relying on an incomplete source led to the discrepancy. Once the appropriate disclosures were identified, the results of the manual review aligned with those of the artefact. This case underscores a key limitation of manual reviews because auditors must devote significant effort to locating and consolidating disclosures that may be dispersed across different reports. In contrast, the artefact's text-mining process systematically extracts the relevant information with consistency and completeness, thereby eliminating the risk of omission.

Table 3. Evaluation of manual and AI-enabled artefact reviews.

Firm	Participant Experience (Years)	Manual Time (Hours)	Manual Result	AI Artefact Result	AI Time (Minutes)
Microsoft	4	3	Compliant (initially)	Flagged (Gap + Plausibility)	~10
Apple	3	4	Compliant	Compliant	~10
Google	5	2	Flagged (Gap + Plausibility)	Flagged (Gap + Plausibility)	~10
Amazon	3	2	Compliant	Compliant	~10
Meta	4	5	Compliant	Compliant	~10

Note: Manual audit results were derived by applying the compliance rules to Scope 2 disclosures in sustainability reports, simulating an auditor's judgment. Manual time was based on practitioner benchmarks [41]. AI time reflects actual runtime for extraction and processing.

Efficiency was assessed by recording the time taken for each manual review. The participants required between two and five hours for each report, depending on the structure and accessibility of disclosures. There was a combined workload of approximately 16 h across the five firms. In contrast, the artefact completed its extraction and analysis for each firm in approximately 10 min and took less than one hour to process the entire dataset. This represents a time saving of more than 90 percent while producing accurate compliance results.

The evaluation demonstrates that the artefact matches the accuracy of manual reviews and eliminates inefficiencies in locating disclosures. The results therefore confirm

the potential of AI-enabled ESG compliance auditing for reducing auditing effort while strengthening reliability. The challenges encountered in manual reviews also help explain why ESG assurance practices can be inconsistent across firms, reinforcing the need for transparent and automated compliance frameworks, such as that developed in this study. The evaluation was limited to sustainability reports from five companies in the technology sector, and the findings are framed as a proof-of-concept in line with the DSR approach.

5. Discussion

The evaluation results confirm that the artefact replicates auditors' judgments while improving the efficiency of an ESG compliance review. AI-enabled automation does not replace auditor judgment but complements the assurance process. By identifying compliance risks rapidly and transparently, the artefact provides auditors and stakeholders with a stronger foundation for evaluating ESG disclosures. Prior research has highlighted the increasing demands placed on assurance services due to the complexity of sustainability reporting and the credibility risks associated with manual verification [5,6]. Professional reports have also documented a shortage of qualified auditors and persistent time pressures for completing engagements, a situation that is being exacerbated by the rapid expansion of ESG reporting requirements [41]. These structural constraints limit the feasibility of manual assurance and raise concerns about consistency across different cases. In this environment, AI-enabled tools offer a practical solution by reducing the time spent on repetitive tasks, improving data capture, and providing standardized compliance checks. This enables auditors to allocate their scarce capacity to high-level judgments while providing stakeholders with reliable sustainability information in a timely manner. The evaluation also contributes to theory by showing how AI can move from interpretive scoring to compliance verification. Prior applications in ESG focused on classification, sentiment analysis, or ratings, which do not test the plausibility of reported values. This study extends the literature by operationalizing the GHG Protocol in a transparent, rule-based artefact built and evaluated through a design science approach.

For auditing practices, the findings suggest that automation can help address the efficiency and consistency challenges faced by ESG assurance professionals. The participants in this study's manual evaluation required between two and five hours per report to locate, extract, and reconcile the relevant disclosures. In contrast, the artefact achieved the same in approximately ten minutes per report, reducing the total time required by more than 90 percent. The scale of this efficiency improvement aligns with the push for auditors to embrace technology to complement their professional judgment. By automating extraction and initial testing, firms can reduce the burden of repetitive work and expand the scope of assurance to a wider set of disclosures. This is particularly relevant as regulators in the United States, European Union, and other jurisdictions advance their mandatory sustainability reporting standards. Under these new frameworks, assurance providers will be expected to verify large volumes of ESG information in a timely and consistent manner. Advanced tools like this study's developed artefact show that audit firms can expand capacity while maintaining quality in this setting. The evaluation also contributes to practice by providing a lightweight tool for different stakeholder groups. Auditors can reduce manual extraction effort and concentrate on higher-risk cases. Regulators can use the tool to focus oversight on firms that report implausible values. Boards can pre-test the credibility of their disclosures, and investors can benchmark reports against transparent thresholds. These applications demonstrate how automation can increase efficiency and enhance comparability while preserving the important role of professional judgment.

The results also have implications for corporate governance and stakeholder oversight. Indeed, boards, investors, and regulators increasingly depend on ESG reports for

decision-making, yet our evaluation revealed that manual reviewers can struggle to locate key disclosures, sometimes leading to misclassifications. This exemplifies how report complexity and data dispersion across multiple sections can pose risks for stakeholders who rely solely on manual reviews. Automated text mining mitigates such risks by capturing all relevant disclosures and subjecting them to consistent testing. This results in a compliance profile that not only flags potential risks but also provides transparent formulas and thresholds. Such transparency enhances trust among stakeholders because they can understand how the results were generated and why certain values were flagged. Moreover, the ability to identify inconsistencies across firms supports the governance objectives of comparability and accountability. For example, the artefact flagged Microsoft and Google for implausible market-based values, raising questions about their reliance on renewable certificates rather than physical decarbonization. Stakeholders benefit from such insights, which may otherwise remain less visible in narrative disclosures.

The evaluation showed that the artefact replicated auditor judgments within the sample and reduced workload by more than 90 percent. These results form the basis of the study's theoretical contribution. Prior literature has reported credibility concerns about sustainability reporting, the resource constraints that auditors face, and the call for more rigorous methods to test disclosures. This research advances the field by operationalizing compliance rules in a functioning artefact that is grounded in DSR. The artefact demonstrates that AI can replicate auditor judgment with precision while addressing structural inefficiencies in manual review. This study offers a methodological contribution by applying artefact design, demonstration, and evaluation to ESG auditing, where applications of AI remain limited. Furthermore, the study extends the dialogue about digital trust technologies. Prior work has shown how blockchain and similar systems can enhance transparency and accountability in reporting [5,6]. The current study complements this by demonstrating how AI can serve as an independent assurance mechanism, thereby increasing the trust in reported values without requiring major shifts in reporting infrastructure.

This discussion, however, emphasizes that the artefact should be viewed as a support tool rather than a replacement for professional assurance. The value of an audit lies not only in detecting inconsistencies but also interpreting results within the broader context of corporate governance, business models, and stakeholder expectations. Automation provides a base layer of standardized checks that in turn allows auditors to focus on areas where professional skepticism and contextual knowledge are invaluable. The artefact's output can also be made accessible to investors, regulators, and other stakeholders who require independent checks before the completion of formal assurance engagements. In this way, AI-enabled tools could democratize access to compliance assessments by reinforcing transparency and reducing information asymmetry. As DSR, the findings are positioned as a proof of concept, with the supporting evidence being limited to five companies in the same industry and one disclosure category. Future research should extend testing to other ESG metrics and industries to assess broader applicability.

6. Conclusions

This study developed and evaluated an AI-enabled artefact for ESG compliance auditing with a focus on Scope 2 GHG disclosures. The artefact extracts reported values, applies compliance rules, and produces transparent outputs. A DSR approach guided the design, demonstration, and evaluation. The artefact was applied to sustainability reports from five companies in the technology sector. The evaluation showed that it replicated auditor judgments within the sample and reduced workload by more than ninety percent.

This study contributes to the literature by showing that AI-enabled artefacts can replicate auditor judgments within the sample and reduce workload by more than 90 percent.

These results demonstrate how DSR can embed compliance rules into artefacts that deliver transparent and replicable outputs. Unlike black-box models, the artefact provides explainable results that align with established assurance procedures. This study reframes the role of AI in ESG auditing from interpretive scoring to compliance verification and presents a proof of concept that aligns methodological rigor with practical relevance. The findings also carry practical implications. For auditors, the artefact offers a scalable response to the challenges of limited capacity, compressed timelines, and expanding ESG disclosure requirements. For other stakeholders, it provides standardized compliance checks that enhance comparability across firms and reduce information asymmetry in sustainability reports.

This study has some limitations, however. The evaluation was restricted to Scope 2 disclosures from five firms in one sector. The results are therefore positioned as a proof of concept consistent with DSR principles. Broader testing across industries and extension to other ESG dimensions, such as Scope 3 emissions, water use, and governance indicators, is necessary to assess the wider robustness. Future research should also examine integration with regulatory frameworks and assess adoption by standard setters.

Overall, this study provides evidence that AI-enabled artefacts can enhance ESG compliance auditing within the technology sector sample. The findings establish a proof-of-concept pathway for greater efficiency, transparency, and reliability in ESG compliance audits by automating repetitive checks while preserving the important role of auditor judgment and governance oversight.

Author Contributions: Conceptualization, E.M.A.; Methodology, E.M.A. and A.M.A.; Validation, E.M.A. and A.M.A.; Formal analysis, E.M.A. and A.M.A.; Investigation, E.M.A. and A.M.A.; Data curation, E.M.A.; Writing—original draft, E.M.A. and A.M.A.; Writing—review & editing, E.M.A. and A.M.A.; Project administration, E.M.A.; Funding acquisition, E.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Open Access program at the American University of Sharjah (Award # OB2604).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The work in this paper was supported, in part, by the Open Access Program from the American University of Sharjah. This paper represents the opinions of the author(s) and does not represent the position or opinions of the American University of Sharjah.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hou, C.; Wen, Y.; Liu, X.; Dong, M. Impacts of regional water shortage information disclosure on public acceptance of recycled water—Evidences from China's urban residents. *J. Clean. Prod.* **2021**, *278*, 123965. [[CrossRef](#)]
2. Michaelowa, A.; Hermwille, L.; Obergassel, W.; Butzengeiger, S. Additionality revisited: Guarding the integrity of market mechanisms under the Paris Agreement. *Clim. Policy* **2019**, *19*, 1211–1224. [[CrossRef](#)]
3. Brander, M.; Gillenwater, M.; Ascui, F. Creative accounting: A critical perspective on the market-based method for reporting purchased electricity (scope 2) emissions. *Energy Policy* **2018**, *112*, 29–33. [[CrossRef](#)]
4. Murray, L. Understanding Large Accounting Firm Perceptions on the Decline of Accounting Graduates and Implications of Supply Shortages. Bachelor's Thesis, University of New Hampshire, Durham, NH, USA, 2024.
5. Alotaibi, E.M.; Issa, H.; Codesso, M. Blockchain-based conceptual model for enhanced transparency in government records: A design science research approach. *Int. J. Inf. Manag. Data Insights* **2025**, *5*, 100304. [[CrossRef](#)]

6. Alotaibi, E.M.; Khallaf, A.; Abdallah, A.A.-N.; Zoubi, T.; Alnesafi, A. Blockchain-driven carbon accountability in supply chains. *Sustainability* **2024**, *16*, 10872. [\[CrossRef\]](#)
7. Bose, S. Evolution of ESG reporting frameworks. In *Values at Work: Sustainable Investing and ESG Reporting*; Springer: New York, NY, USA, 2020; pp. 13–33. [\[CrossRef\]](#)
8. Carungu, J.; Dimes, R.; Molinari, M. EFRAG and ISSB: Tensions and opportunities for convergence in the quest for the standardisation of sustainability reporting standards. *Manag. Decis.* **2025**, *ahead-of-print*. [\[CrossRef\]](#)
9. Hummel, K.; Jobst, D. An overview of corporate sustainability reporting legislation in the European Union. *Account. Eur.* **2024**, *21*, 320–355. [\[CrossRef\]](#)
10. WRI. *Greenhouse Gas Protocol: Scope 2*; World Resources Institute: Washington, DC, USA, 2015.
11. Kreibich, N.; Hermwille, L. Caught in between: Credibility and feasibility of the voluntary carbon market post-2020. *Clim. Policy* **2021**, *21*, 939–957. [\[CrossRef\]](#)
12. Soares, I.V.; Yarime, M.; Klemun, M.M. Estimating GHG emissions from cloud computing: Sources of inaccuracy, opportunities and challenges in location-based and use-based approaches. *Clim. Policy* **2025**, 1–19. [\[CrossRef\]](#)
13. Perego, P.; Kolk, A. Multinationals' accountability on sustainability: The evolution of third-party assurance of sustainability reports. *J. Bus. Ethics* **2012**, *110*, 173–190. [\[CrossRef\]](#)
14. Boiral, O.; Heras-Saizarbitoria, I. Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? *J. Clean. Prod.* **2020**, *243*, 118596. [\[CrossRef\]](#)
15. Christensen, H.B.; Hail, L.; Leuz, C. Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Rev. Account. Stud.* **2021**, *26*, 1176–1248. [\[CrossRef\]](#)
16. Moroney, R.; Trotman, K.T. Differences in auditors' materiality assessments when auditing financial statements and sustainability reports. *Contemp. Account. Res.* **2016**, *33*, 551–575. [\[CrossRef\]](#)
17. Farooq, M.B.; de Villiers, C. The shaping of sustainability assurance through the competition between accounting and non-accounting providers. *Account. Audit. Account. J.* **2019**, *32*, 307–336. [\[CrossRef\]](#)
18. Casey, R.J.; Grenier, J.H. Understanding and contributing to the enigma of corporate social responsibility (CSR) assurance in the United States. *Audit. A J. Pract. Theory* **2015**, *34*, 97–130. [\[CrossRef\]](#)
19. ISSB. *IFRS S1: Sustainability Disclosure Standard*; International Sustainability Standards Board: London, UK, June 2023. Available online: <https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards-issb/english/2023/issued/part-a/issb-2023-a-ifrs-s1-general-requirements-for-disclosure-of-sustainability-related-financial-information.pdf> (accessed on 20 October 2025).
20. ISAE 3000; Assurance Engagements Other Than Audits or Reviews of Historical Financial Information. International Auditing and Assurance Standards Board, International Federation of Accountants: New York, NY, USA, 2013.
21. Nangoy, G.F.; Meiden, C. Analysis Of The Quality of Isae 3000 Assurance Statements On Corporate Sustainability Reports Across Stock Exchanges In Several Countries: A 2020–2022 Study. *Int. J. Soc. Sci.* **2024**, *3*, 581–588. [\[CrossRef\]](#)
22. Li, N.; Kim, M.; Dai, J.; Vasarhelyi, M.A. Using artificial intelligence in ESG assurance. *J. Emerg. Technol. Account.* **2024**, *21*, 83–99. [\[CrossRef\]](#)
23. Brown-Liburd, H.; Issa, H.; Lombardi, D. Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Account. Horiz.* **2015**, *29*, 451–468. [\[CrossRef\]](#)
24. Appelbaum, D.; Kogan, A.; Vasarhelyi, M.; Yan, Z. Impact of business analytics and enterprise systems on managerial accounting. *Int. J. Account. Inf. Syst.* **2017**, *25*, 29–44. [\[CrossRef\]](#)
25. Kokina, J.; Davenport, T.H. The emergence of artificial intelligence: How automation is changing auditing. *J. Emerg. Technol. Account.* **2017**, *14*, 115–122. [\[CrossRef\]](#)
26. Smeuninx, N.; De Clerck, B.; Aerts, W. Measuring the readability of sustainability reports: A corpus-based analysis through standard formulae and NLP. *Int. J. Bus. Commun.* **2020**, *57*, 52–85. [\[CrossRef\]](#)
27. Mohamed Riyath, M.I.; Inun Jariya, A.M. The role of ESG reporting, artificial intelligence, stakeholders and innovation performance in fostering sustainability culture and climate resilience. *J. Financ. Report. Account.* **2024**. [\[CrossRef\]](#)
28. Amel-Zadeh, A.; Serafeim, G. Why and how investors use ESG information: Evidence from a global survey. *Financ. Anal. J.* **2018**, *74*, 87–103. [\[CrossRef\]](#)
29. Saberi, S.; Kouhizadeh, M.; Sarkis, J.; Shen, L. Blockchain technology and its relationships to sustainable supply chain management. *Int. J. Prod. Res.* **2019**, *57*, 2117–2135. [\[CrossRef\]](#)
30. Casino, F.; Dasaklis, T.K.; Patsakis, C. A systematic literature review of blockchain-based applications: Current status, classification and open issues. *Telemat. Inform.* **2019**, *36*, 55–81. [\[CrossRef\]](#)
31. Khan, M.; Serafeim, G.; Yoon, A. Corporate sustainability: First evidence on materiality. *Account. Rev.* **2016**, *91*, 1697–1724. [\[CrossRef\]](#)
32. Ostojic, S.; Backes, J.G.; Kowalski, M.; Traverso, M. Beyond Compliance: A Deep Dive Into Improving Sustainability Reporting Quality with LCSA Indicators. *Standards* **2024**, *4*, 196–246. [\[CrossRef\]](#)

33. Sewpersadh, N.S. Adaptive structural audit processes as shaped by emerging technologies. *Int. J. Account. Inf. Syst.* **2025**, *56*, 100735. [\[CrossRef\]](#)
34. Lombardi, D.; Brown-Liburd, H.L.; Munoko, I. Using an Interactive Artificial Intelligence System to Augment Auditor Judgment in a Complex Task. *SSRN Electron. J.* **2023**, 4318689. [\[CrossRef\]](#)
35. Wu, J.-Y.; Nataraj, V.; Day, M.-Y. Generative AI in ESG Reporting: A Systematic Review. In Proceedings of the International Conference on Advances in Social Networks Analysis and Mining, University of Calabria, Rende (CS), Calabria, Italy, 2–5 September 2024; pp. 337–352.
36. Gregor, S.; Hevner, A.R. Positioning and presenting design science research for maximum impact. *Manag. Inf. Syst. Quarterly* **2013**, *37*, 337–355. [\[CrossRef\]](#)
37. Mahapatra, A.; Srivastava, N.; Srivastava, J. Contextual anomaly detection in text data. *Algorithms* **2012**, *5*, 469–489. [\[CrossRef\]](#)
38. Raadal, H.L.; Ghorbani, M. *Residual Mix Methodology for I-REC Issuing Countries*; Norwegian Institute for Sustainability Research: Fredrikstad, Norway, 2023.
39. OECD. *OECD Climate Change and Corporate Governance, Corporate Governance*; OECD Publishing: Paris, France, 2022.
40. Zhou, R.; Lou, J.; He, B. Greening corporate environmental, social, and governance performance: The impact of China's carbon emissions trading pilot policy on listed companies. *Sustainability* **2025**, *17*, 963. [\[CrossRef\]](#)
41. IFAC. *The State of Play in Sustainability Assurance*; International Federation of Accountants (IFAC): New York, NY, USA; AICPA & CIMA: Durham, NC, USA, 2023.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.