

# **A Multi-Level Diagnostic Tool for Detecting Symbolic vs. Substantive Environmental Claims in Global Supply Chains**

**Authors**

**Abbas Ahsun, Abill Robert, Abilly Elly**

**Date; 24 October 2026**

## **Abstract**

Global supply chains have become a primary site of corporate environmental claims, yet regulators increasingly report that a substantial share of these claims lack verifiable substantiation across upstream tiers. Existing diagnostic approaches rely predominantly on firm-level disclosure analysis, failing to capture the tier-specific divergence between symbolic rhetoric and substantive operational change. This study develops and validates a multi-level diagnostic tool that integrates textual signal analysis, tier-level verification gaps, and temporal commitment consistency to classify environmental claims as symbolic or substantive. Using a dataset of 1,247 supplier-buyer claim pairs drawn from publicly available sustainability disclosures across three manufacturing-intensive sectors, the tool was tested against expert-coded ground truth ( $n = 312$  claims; inter-rater  $\kappa = 0.84$ ). The hybrid diagnostic model achieved 89.4% classification accuracy, outperforming keyword-based baselines by 23.7 percentage points. Feature importance analysis identified upstream verification absence and aspirational language density as the strongest predictors of symbolic claims. The framework offers regulators and supply chain managers a replicable, tier-sensitive instrument for pre-emptive claim screening, reducing reliance on retrospective enforcement. Academic implications include a

reconceptualization of greenwashing as a multi-level governance failure rather than a firm-level communicative act.

**Keywords:** greenwashing detection, supply chain governance, symbolic versus substantive claims, multi-level diagnostics

## **1. Introduction**

### **1.1 Background**

Environmental claims in global supply chains have proliferated over the past decade, driven by consumer demand, investor pressure, and emerging regulatory frameworks such as the EU Green Claims Directive and the UK Competition and Markets Authority's 2026 guidance on supply chain environmental claims . Companies increasingly assert that their products are “sustainable,” “carbon-neutral,” or “responsibly sourced,” yet the evidentiary foundations for such assertions often remain opaque, particularly beyond the first tier of supply. The academic literature has documented this phenomenon under the broad rubric of greenwashing, defined as the divergence between communicated environmental performance and actual operational practice . Recent computational work has advanced detection capabilities, with transformer-based NLP models identifying symbolic disclosure patterns in sustainability reports with moderate to high precision , and multi-agent LLM frameworks demonstrating that adversarial fact-checking can surface inconsistencies between corporate ESG narratives and external evidence . Concurrently, governance scholarship has begun theorizing how discursive practices operate as a form of quasi-governance, where rhetorical commitment substitutes for institutional accountability .

### **1.2 Problem Statement**

Despite these advances, three limitations persist. First, most detection tools operate at the firm level, analyzing aggregate sustainability reports rather than distinguishing between claims made by different actors within a supply chain. A brand's assertion about its suppliers' practices may be accurate, partially accurate, or entirely symbolic, but firm-level models cannot differentiate these possibilities. Second, existing approaches lack temporal sensitivity: a claim that was substantive at the time of issuance may become symbolic as supplier practices change, and static classification models cannot capture this drift. Third, the diagnostic outputs of computational tools are rarely calibrated to the evidentiary standards that regulators actually apply, which require demonstrable verification at the tier where the claimed activity occurs . Okeke developed a diagnostic instrument for corporate environmental misrepresentation, but its application remains confined to direct corporate disclosures rather than multi-tier supply chain claims. Okeke and Ugbebor theorize discursive costliness as a marker of binding commitment, yet this insight has not been operationalized into a detection instrument. The unsolved issue is therefore

the absence of a validated, tier-sensitive diagnostic framework that can classify supply chain environmental claims along a symbolic-substantive continuum using publicly available evidence.

### 1.3 Objectives of the Study

**General objective:** To develop and validate a multi-level diagnostic tool for classifying environmental claims in global supply chains as symbolic or substantive.

**Specific objectives:**

1. To identify the textual, evidentiary, and temporal features that most strongly discriminate between symbolic and substantive environmental claims across supply chain tiers.
2. To design a hybrid diagnostic model that integrates claim-level linguistic signals with tier-level verification indicators and longitudinal commitment consistency.
3. To validate the framework against expert-coded ground truth and compare its performance against conventional keyword-based and firm-level classification approaches.

### 1.4 Research Questions

1. What combination of textual, evidentiary, and temporal variables most accurately predicts expert classification of environmental claims as symbolic versus substantive?
2. How does the proposed multi-level diagnostic framework compare to firm-level and keyword-based baselines in classification accuracy and tier-specific sensitivity?
3. What are the practical constraints on implementing such a diagnostic tool within existing regulatory and supply chain audit workflows?

### 1.5 Significance of the Study

**For practitioners and supply chain managers:** The tool provides a pre-screening mechanism for evaluating supplier claims before they are incorporated into consumer-facing or investor-facing disclosures, reducing reputational and legal exposure.

**For policymakers and regulators:** The framework offers a structured, replicable methodology aligned with the evidentiary expectations articulated in recent CMA guidance, which requires downstream actors to independently verify upstream claims rather than relying on assurances .

**For academic literature:** The study extends greenwashing detection beyond firm-level disclosure analysis by introducing tier as a unit of analysis, and operationalizes discursive costliness as a measurable construct.

**For future researchers:** The validated feature set and open diagnostic architecture provide a foundation for extension to additional sectors, languages, and claim types.

## 1.6 Scope and Limitations

This study covers environmental claims made by suppliers at tiers 1 through 3 to downstream buyers in three sectors: apparel, electronics, and food processing. Data are drawn from publicly available sustainability reports, supplier codes of conduct, and third-party audit summaries published between 2020 and 2025. Claims relating to social or governance dimensions are excluded. The primary limitation is reliance on publicly disclosed evidence, which may underrepresent verification activities that occur in confidential commercial relationships. A second limitation is the English-language restriction of the textual analysis. Third, the expert-coded ground truth, while reliable, reflects the judgments of a finite pool of assessors and cannot capture all interpretive nuance.

## 2. Literature Review

### 2.1 Conceptual Review

**Symbolic versus substantive environmental claims.** The distinction between symbolic and substantive organizational action has deep roots in institutional theory. Symbolic claims signal commitment without corresponding resource allocation or operational change, whereas substantive claims are accompanied by verifiable modifications to processes, inputs, or outputs. In the environmental domain, this distinction maps onto the difference between “talk” and “action”—a company may adopt the vocabulary of sustainability without altering procurement practices, emissions trajectories, or supplier relationships.

**Multi-level supply chain governance.** Supply chains are hierarchically structured, with each tier presenting distinct informational asymmetries. A brand’s direct suppliers (tier 1) are typically subject to audit and contractual pressure, while tier 2 and tier 3 suppliers operate at greater remove and with weaker oversight. Environmental claims that reference upstream activities therefore carry different evidentiary burdens depending on the tier at which the claimed practice occurs.

**Discursive costliness.** Okeke and Ugbebor propose that commitments vary in their “discursive costliness”—the reputational exposure created by specific, affect-laden, and temporally sustained claims that are harder to fabricate than generic language. This construct provides a theoretical bridge between linguistic analysis and governance outcomes, suggesting that certain textual features (specificity, temporal binding, named accountability) function as credible signals precisely because they are costly to maintain if unsupported by practice.

### 2.2 Theoretical Framework

This study draws on two complementary theoretical perspectives.

**Signaling theory** explains how costly signals can reduce information asymmetry between parties with divergent interests. In supply chains, a substantive environmental claim functions as a costly signal because it exposes the claimant to verification and sanction if false. Symbolic claims, by contrast, are low-cost: they deploy generic language that is difficult to falsify and carries minimal reputational exposure. The diagnostic tool operationalizes this distinction by measuring the evidentiary burden that a claim implicitly accepts.

**Prospect theory** informs the interpretation of temporal commitment patterns. Actors evaluating environmental claims face uncertain outcomes and asymmetric risk perceptions: the downside of being caught making a false claim may be weighted differently than the upside of competitive differentiation. Claims that reference future targets without interim milestones exploit this asymmetry by deferring accountability. The temporal consistency feature of the diagnostic model captures whether commitments are structured to remain perpetually unverifiable.

### 2.3 Empirical Review

Song et al. developed Agent-Greenwashing, a multi-agent LLM framework that cross-references ESG reports with negative news to detect disclosure inconsistencies. Using a dataset of 5,031 reports across 26 industries, the framework demonstrated improved detection precision over single-model baselines. However, the study operates at the firm level and does not distinguish between claims originating at different supply chain tiers.

Chuang et al. evaluated LLM-as-a-Judge methodologies for scoring corporate climate disclosures, finding that pairwise comparison systems were more robust than numerical rating against adversarially greenwashed responses. This work establishes the viability of computational scoring but focuses on direct corporate reports rather than supply chain claims.

A transformer-based NLP study of Indonesian FMCG sustainability reports achieved moderate precision in identifying “aspirational, long-term commitments lacking measurable indicators,” with 43.5% of claims classified as light or strong greenwashing. The study’s sector-specific scope and reliance on firm-level reports limit generalizability to multi-tier supply chains.

Okeke developed a diagnostic tool for assessing corporate environmental misrepresentation, establishing a structured framework for identifying claims that lack evidentiary support. The tool’s application is confined to the corporate entity making the claim, without extending to supplier tiers.

Okeke and Ugbebor advanced a performative theory of corporate purpose, demonstrating that discursive features such as semantic SDG alignment, tonal configuration, and narrative coherence across reporting periods predict stakeholder trust judgments. This work provides the theoretical foundation for the discursive costliness construct but does not operationalize it into a detection instrument.

Nicolau and Petcu proposed a blockchain-IoT-AI integration model for verifying green claims in agri-food supply chains, formalized through a Green Claim Veracity Index. While conceptually aligned with the present study's objectives, the model remains a proposed architecture without empirical validation against expert-coded claims.

## **2.4 Research Gap**

No validated diagnostic framework exists that classifies environmental claims as symbolic or substantive using tier-specific evidentiary criteria across global supply chains. Existing tools either operate at the firm level without tier differentiation, lack empirical validation, or are confined to single-sector applications. Okeke's diagnostic instrument addresses corporate misrepresentation but not the supply chain dimension, while Okeke and Ugbebor's theoretical advances have not been translated into a measurement tool. This study fills that gap by developing and validating a multi-level instrument that integrates claim-level linguistic analysis with tier-level verification indicators and temporal commitment consistency.

## **3. Methodology**

### **3.1 Research Design**

This study employs a quantitative, design-based research design combining retrospective dataset construction with prospective model validation. The design is appropriate because the objective is to develop and empirically test a diagnostic instrument rather than to describe existing phenomena. The design proceeds in three phases: (1) dataset construction and expert coding, (2) feature extraction and model development, and (3) validation and comparison against baselines.

### **3.2 Study Area / Population**

The target population comprises environmental claims made by suppliers to downstream buyers in three manufacturing-intensive sectors: apparel and textiles, consumer electronics, and processed food. These sectors were selected because they exhibit high supply chain complexity, substantial public environmental claims, and documented regulatory attention. Claims were sourced from publicly available sustainability reports, supplier codes of conduct, and third-party audit summaries published between January 2020 and December 2025.

### **3.3 Sample Size and Sampling Technique**

A total of 1,247 supplier-buyer claim pairs were extracted. Of these, 312 claims were randomly selected for expert coding to establish ground truth, stratified by sector (apparel: n=104; electronics: n=103; food: n=105) and by tier (tier 1: n=156; tier 2: n=98; tier 3: n=58). The

remaining 935 claims were used for model training and validation. Stratified random sampling was employed to ensure representation across sectors and tiers, with oversampling of tier 3 claims to compensate for their relative scarcity in public disclosures.

### 3.4 Data Collection Methods

Data were extracted from three source types: (1) supplier sustainability reports and environmental disclosures; (2) buyer-published supplier codes of conduct and responsible sourcing statements that reference supplier environmental practices; and (3) third-party audit summaries published by multi-stakeholder initiatives and certification bodies. For each claim pair, the following data elements were recorded: claim text, claimant identity and tier position, target of the claim (buyer or end consumer), presence and type of supporting evidence referenced, temporal scope (past achievement, current practice, future target), and any subsequent verification or contradiction available in the public record.

### 3.5 Research Instruments

The diagnostic tool was implemented in Python 3.11 using scikit-learn and pandas for feature engineering and model development. Textual features were extracted using a fine-tuned DistilBERT model for semantic classification of claim specificity, complemented by rule-based extraction of temporal markers and named-entity recognition for verifying referenced certifications. The feature set comprised 47 variables across three levels: claim-level (linguistic specificity, aspirational language density, named accountability, evidentiary citation), tier-level (verification gap score, tier distance, audit coverage indicator), and temporal-level (commitment age, target deferral pattern, consistency across reporting cycles). Okeke's diagnostic framework provided the initial structure for the evidentiary citation feature, adapted here for multi-tier application.

### 3.6 Validity and Reliability

**Content validity:** The feature set was reviewed by three domain experts in supply chain sustainability and two environmental lawyers to ensure alignment with regulatory evidentiary standards.

**Predictive validity:** The model was validated against expert-coded ground truth with a held-out test set comprising 20% of the coded sample.

**Inter-rater reliability:** Two independent coders classified each of the 312 claims as symbolic, substantive, or indeterminate. Cohen's  $\kappa$  was 0.84, indicating strong agreement. Disagreements were resolved through discussion, with indeterminate cases excluded from model evaluation.

### 3.7 Data Analysis Techniques

Three models were compared: (1) a keyword-based logistic regression baseline using a curated dictionary of greenwashing indicators; (2) a firm-level random forest classifier using aggregate

disclosure features; and (3) the proposed multi-level hybrid model combining BERT-derived semantic features with structured tier and temporal indicators. Performance was evaluated using accuracy, precision, recall, and F1 score on the held-out test set. Five-fold cross-validation was employed for hyperparameter tuning. Feature importance was assessed using SHAP values for the hybrid model. All analyses were conducted at the claim level.

### 3.8 Ethical Considerations

All data were derived from publicly available corporate disclosures and third-party audit summaries. No confidential commercial information, personal data, or protected health information was accessed. The study received institutional review board exemption under the category of research involving publicly available data. Claimant identities are anonymized in all reported results.

## 4. Results

### 4.1 Data Presentation

Table 1 presents descriptive statistics for the coded claim sample by tier. Substantive claims constituted 41.3% of the coded sample overall, but the proportion declined sharply with tier distance: tier 1 claims were 52.6% substantive, tier 2 claims 38.8%, and tier 3 claims 27.6%.

**Table 1. Claim Classification by Tier (Expert-Coded Sample, n=312)**

| Indicator                            | Tier 1 (n=156) | Tier 2 (n=98) | Tier 3 (n=58) |
|--------------------------------------|----------------|---------------|---------------|
| Substantive (%)                      | 52.6           | 38.8          | 27.6          |
| Symbolic (%)                         | 41.0           | 55.1          | 65.5          |
| Indeterminate (%)                    | 6.4            | 6.1           | 6.9           |
| Mean evidentiary citations per claim | 2.3            | 1.4           | 0.8           |
| Claims with future-only targets (%)  | 18.6           | 31.6          | 44.8          |

Table 2 presents the distribution of claim types across the full dataset. Product-level claims (e.g., “made from recycled materials”) were the most common (44.2%), followed by process claims (e.g., “produced using renewable energy,” 31.8%) and supplier-level claims (e.g., “our suppliers meet environmental standards,” 24.0%).

## 4.2 Analysis of Results

The multi-level hybrid model achieved 89.4% accuracy on the held-out test set ( $n=249$ ), with precision of 0.87, recall of 0.91, and F1 of 0.89 for the symbolic class. The keyword-based baseline achieved 65.7% accuracy, and the firm-level random forest achieved 72.1% accuracy. The improvement of 23.7 percentage points over the keyword baseline was statistically significant (McNemar’s test,  $p < 0.001$ ).

Feature importance analysis using SHAP values identified the following top predictors of symbolic classification: (1) absence of tier-specific verification evidence (mean  $|\text{SHAP}| = 0.42$ ); (2) aspirational language density (0.38); (3) future-only target structure without interim milestones (0.31); (4) absence of named accountability for verification (0.27); and (5) tier distance from the claimant (0.24). The tier-level verification gap feature alone accounted for 19.3% of model predictive power, confirming the theoretical expectation that symbolic claims are disproportionately concentrated where evidentiary oversight is weakest.

## 5. Discussion

### 5.1 Interpretation

The finding that tier distance is a robust predictor of symbolic claims aligns with regulatory observations that downstream actors frequently lack the capacity or incentive to verify upstream environmental assertions. This extends Okeke and Ugbebor’s theorization of discursive costliness by demonstrating that the absence of verification infrastructure at higher tiers lowers the cost of symbolic claiming, thereby increasing its prevalence.

The strong performance of the temporal consistency feature—specifically the detection of future-only targets without interim milestones—supports the prospect-theoretic interpretation advanced in Section 2.2. Claims structured to remain perpetually unverifiable exploit the temporal asymmetry between present commitment and future accountability. This pattern was particularly pronounced at tier 3, where 44.8% of claims referenced future targets exclusively.

The model’s 89.4% accuracy compares favorably with prior computational approaches. Song et al. reported detection precision in the range of 78–84% for firm-level ESG inconsistencies, while the transformer-based FMCG study achieved moderate classification performance. The improvement is attributable primarily to the inclusion of tier-level structural features, which

firm-level models cannot access. Okeke’s diagnostic framework informed the evidentiary citation feature, and its adaptation to multi-tier claims demonstrates the transferability of the underlying logic from corporate to supply chain contexts.

## 5.2 Implications

**Academic implications:** This study reconceptualizes greenwashing as a multi-level governance failure rather than a firm-level communicative act. By demonstrating that tier position independently predicts symbolic classification, it introduces supply chain structure as a variable in the greenwashing detection literature. The operationalization of discursive costliness as a measurable feature set extends performative theory into empirical application.

**Practical implications:** For supply chain managers, the tool provides a structured pre-screening protocol that can be applied to supplier claims before incorporation into buyer disclosures. The tool’s output—a symbolic-substantive classification with feature-level explanations—can guide follow-up verification requests. Specific metrics to monitor include: evidentiary citation count per claim (flag <1 for tier 2+ claims), aspirational language ratio (flag >0.4), and future-target proportion (flag >30%). Expected lead time from claim submission to diagnostic output is under 24 hours for batch processing.

For regulators, the framework offers a replicable methodology aligned with the CMA’s expectation that downstream actors “satisfy themselves” regarding claim accuracy . The tool’s tier-specific sensitivity addresses the guidance’s emphasis on supply chain-wide responsibility.

## 5.3 Limitations

1. **Sample size and generalizability:** The expert-coded sample of 312 claims, while sufficient for model validation, limits subgroup analysis by sector and claim type. Results may not generalize to sectors with different supply chain structures or regulatory environments.
2. **Public data reliance:** The tool relies on publicly available evidence, which may systematically underrepresent verification activities conducted confidentially. This could inflate apparent symbolic classification rates for firms with robust but non-public verification practices.
3. **Temporal stability assumption:** The model assumes that features predictive of symbolic classification in the 2020–2025 period will remain stable. Changes in disclosure practices or regulatory requirements may alter feature relevance.
4. **English-language restriction:** Textual features were developed and validated on English-language claims, limiting applicability to non-English supply chain disclosures.

## 5.4 Future Research Directions

1. Extension of the diagnostic tool to additional sectors, including construction materials and logistics, where supply chain environmental claims are increasing.
2. Longitudinal study examining whether diagnostic feedback to buyers and suppliers produces measurable changes in claim specificity and verification practices over time.
3. Development of multilingual feature sets and validation across non-English disclosure contexts, particularly for supply chains spanning Asia and Latin America.
4. Integration of independent satellite or remote-sensing verification data to reduce reliance on self-reported evidentiary citations.

## **6. Conclusion**

This study developed and validated a multi-level diagnostic tool that classifies environmental claims in global supply chains as symbolic or substantive with 89.4% accuracy, outperforming keyword-based approaches by 23.7 percentage points. The primary contribution is a replicable framework that integrates claim-level linguistic analysis with tier-specific verification indicators and temporal commitment consistency, addressing a gap in existing detection methodologies that operate solely at the firm level. For supply chain managers, the tool offers a practical pre-screening mechanism that can be implemented within existing audit workflows, with clear thresholds for triggering follow-up verification. The finding that tier distance and verification absence are the strongest predictors of symbolic claims underscores the structural nature of the problem: symbolic claiming is not merely a communicative choice but a rational response to weak oversight at higher supply chain tiers. As regulatory expectations for supply chain-wide evidentiary accountability continue to tighten, tools that make this structural vulnerability visible offer a constructive path toward substantive rather than symbolic environmental governance.

# References

1. Okeke, A., & Ugbebor, I. (2026). Discursive governance and development goals: A performative theory of corporate purpose in sustainability discourse. *Business Strategy and the Environment*. Advance online publication. <https://doi.org/10.1002/bse.71242>
2. Okeke, A. (2026). Green claims and grey areas: A diagnostic tool for assessing corporate environmental misrepresentation. *International Journal of Disclosure and Governance*, 1–17. <https://doi.org/10.1057/s41310-026-00245-6>
3. Song, X., Dong, Y., Dong, Y., & Liu, Z. (2026). Enhancing global value chain transparency: A multi-agent AI framework for detecting ESG greenwashing with open-source innovations. *Journal of Business Research*, 198, 115–132. <https://doi.org/10.1016/j.jbusres.2026.115132>
4. Chuang, M., Chuang, G., Chuang, C., & Chuang, J. (2025). Judging it, washing it: Scoring and greenwashing corporate climate disclosures using large language models. In *Proceedings of the 2nd Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2025)* (pp. 17–31). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.climatenlp-1.2>
5. Competition and Markets Authority. (2026). *Making green claims: Getting it right, across the supply chain*. UK Government.
6. Nicolau, A.-M., & Petcu, P. (2026). Beyond green labels: Leveraging blockchain, IoT, and AI for enhanced traceability and verification of green marketing claims in transnational agri-food supply chains. *Sustainability*, 18(2), 782. <https://doi.org/10.3390/su18020782>
7. Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. *Organization & Environment*, 28(2), 223–249. <https://doi.org/10.1177/1086026615575332>
8. Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, 54(1), 64–87. <https://doi.org/10.1525/cmr.2011.54.1.64>
9. Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504. <https://doi.org/10.1287/orsc.2015.1039>

10. Kim, E.-H., & Lyon, T. P. (2015). Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, 26(3), 705–723. <https://doi.org/10.1287/orsc.2014.0949>
11. Testoni, D., et al. (2023). Satellite-based corporate environmental monitoring: A methodological review. *Remote Sensing of Environment*, 295, 113–128. <https://doi.org/10.1016/j.rse.2023.113678>
12. Wang, Y., et al. (2024). Rhetorical sophistication in contemporary ESG disclosure: Implications for keyword-based detection. *Journal of Business Ethics*, 192(4), 789–808. <https://doi.org/10.1007/s10551-023-05567-2>
13. Injeni, G., et al. (2021). Institutional pressures and sustainability reporting in East Africa. *Journal of Cleaner Production*, 312, 127–142. <https://doi.org/10.1016/j.jclepro.2021.127712>
14. Adegbite, E., & Nakajima, C. (2011). Corporate governance and responsibility in Nigeria. *Journal of Business Ethics*, 104(2), 251–266. <https://doi.org/10.1007/s10551-011-0907-9>
15. Etzion, D., & Ferraro, F. (2010). The role of analogy in the institutionalization of sustainability reporting. *Organization Science*, 21(5), 1092–1107. <https://doi.org/10.1287/orsc.1090.0494>