The Authenticity Feedback Model (AFM): Quantifying Feedback Loops in Perception–Reality Gaps within Architectural Sustainability and ESG Legitimacy

By Thejan Anil

Affiliation:

Independent Researcher – Sustainable Strategy and ESG Communication

Contact:

thejananil@gmail.com

Index

The Authenticity Feedback Model (AFM): Quantifying Feedback Loops in Perception–Reality Gaps within Architectural Sustain-Sustainability and ESG Legitimacy

1. Abstract

The Authenticity Feedback Model (AFM) presents a cyclical framework for evaluating sustainability and ethical integrity in architectural projects. Developed to address the persistent disconnect between ESG rhetoric and operational reality, AFM transforms environmental assessment from a static certification process into a dynamic, data-driven feedback system. The model integrates six analytical layers—contextual fairness, rhetoric, translation, perception, reality, and verification—to capture how intent evolves into measurable outcomes. By quantifying authenticity through iterative comparisons between stated benchmarks and observed performance, AFM identifies and corrects perception–reality gaps. Its built-in anomaly detection and longitudinal feedback loop enable continuous recalibration across project cycles. While conceptual in scope, AFM establishes a scalable foundation for empirical validation and cross-disciplinary adaptation within sustainable architecture and ESG governance.

2. Introduction

ESG has become a central pillar of contemporary corporate strategy. Over the past few decades, narratives surrounding ESG have evolved into the fulcrum of board-level decision-making across industries and geographies. Yet this prominence has also created opportunities for organizations to deploy ESG discourse as strategic theatrics—employing the language of environmental and social responsibility as a façade for profit-oriented objectives. In such cases, stakeholders and consumers often become victims of these rhetorical performances, struggling to distinguish genuine commitment from opportunistic branding—a phenomenon widely recognized as greenwashing.

This legitimacy crisis arises from systemic weaknesses such as information asymmetry, signal inflation, and the absence of credible empirical validation. Corporate sustainability communications—through press releases, integrated reports, and promotional narratives—

often rely on rhetorical framing rather than verifiable performance data, distorting stakeholder perception and amplifying symbolic compliance. Numerous interdisciplinary studies have evidenced this dissonance, forming the conceptual impetus behind the development of the Authenticity Feedback Model (AFM). The model aims to mitigate performative ESG strategies by introducing a cyclical, data-anchored feedback mechanism that restores transparency and accountability in corporate communication.

While several global frameworks—such as LEED and BREEAM—have advanced sustainability assessment, their efficacy diminishes at the performance stage, where static certification fails to reflect dynamic operational realities. AFM addresses this gap through a cyclic architecture that continuously integrates post-performance metrics, stakeholder feedback, and communicative analysis. By establishing recurrent evaluation checkpoints, the model enables adaptive learning and systemic evolution in response to policy shifts and governance transitions. Unlike conventional post-certification validation, AFM combines empirical data with linguistic assessment, analyzing both operational performance and the rhetorical tone embedded in ESG narratives.

The model is structured into five stages, each grounded in theoretical insights drawn from Sen (1992, 1999), Spence (1973), Meyer and Rowan (1977), and Wiener (1948). These works collectively inform the ethical, communicative, institutional, and systemic dimensions of AFM. The framework aspires to function not merely as a diagnostic tool but as an ethical infrastructure—bridging moral philosophy and empirical governance. Its implementation could transform sustainability assessment from a rhetorical exercise into an instrument of genuine accountability, ensuring that corporate communication aligns with lived environmental and social realities. In essence, AFM serves as both a corrective mechanism and a moral compass—guiding firms toward authentic, verifiable, and equitable ESG engagement.

Once the framework is integrated into organizational and regulatory systems, it can significantly enhance the long-term performance and transparency of the built-environment sector. By embedding continuous feedback into sustainability evaluation, AFM strengthens the credibility of ESG reporting and fosters genuine ecological and social progress—reducing the dependence on rhetorical communication as a substitute for measurable action.

3. Literature Review

Machura-Urbániak (2024) critically examines the real-estate sector as a major contributor to environmental degradation—not only through its intensive resource consumption and carbon emissions but also through its reliance on misleading sustainability communication (Machura-Urbániak, 2024). The paper argues that firms within this sector frequently employ marketing rhetoric, certification labels, and symbolic language to project an illusion of ecological responsibility that lacks empirical substance. Terms such as natural or prefixes like *eco-* and bio- are deliberately chosen to evoke environmental authenticity and emotionally appeal to sustainability-minded consumers. This linguistic framing, which the author defines as semantic greenwashing, enables corporations to maintain reputational legitimacy and stakeholder approval while enacting minimal substantive environmental reform. Supporting evidence from the European Commission reveals that approximately 53 percent of projects marketed as "eco" initiatives are based on vague or unverifiable claims—a trend that erodes consumer trust and weakens long-term corporate credibility (European Commission, 2021, as cited in Machura-Urbániak, 2024).

The study situates its analysis within the Luxembourg real-estate landscape, a country that balances sustainable innovation with the preservation of its medieval architectural heritage. Drawing on data from the International Energy Agency (IEA, 2021), Machura-Urbániak underscores that buildings account for nearly 30 percent of global energy consumption and 27 percent of CO_2 emissions, reinforcing the urgent need for decarbonization across both new and heritage structures. The paper emphasizes that sustainability in real estate must extend beyond design aesthetics and include verifiable performance metrics integrated into ESG frameworks (IEA, 2021; Machura-Urbániak, 2024).

A major contribution of the study lies in its analysis of the Sustainable Finance Disclosure Regulation (SFDR) as a regulatory mechanism to counteract symbolic compliance. By mandating the disclosure of measurable and verifiable sustainability information, SFDR enhances transparency and equips investors to evaluate the authenticity of ESG claims (Machura-Urbániak, 2024). In doing so, it directly mitigates the risks of greenwashing that stem from ambiguous or purely reputational sustainability narratives. Transparency thus emerges as a central theme of the study—aligning with the European Union's broader commitment to steering the sector toward credible, net-zero transitions.

In conclusion, Machura-Urbániak (2024) demonstrates that ESG narratives in real estate often evolve into strategic communication tools shaped by consumer psychology, regulatory pressure, and competitive signalling. The paper argues that authentic disclosure and consistent verification are essential if sustainability rhetoric is to yield real environmental impact. This conceptual intersection—between communication authenticity, regulatory accountability, and reputational legitimacy—provides a theoretical bridge to the development of Authenticity-Feedback Models (AFM) such as the one advanced in the present study.

The conceptual foundation laid by Machura-Urbániak (2024) establishes how semantic manipulation and regulatory opacity foster reputational sustainability without verifiable reform. To extend this argument, it is essential to examine the deeper organizational and behavioural psychology behind such communication. Delmas and Burbano (2011) provide that theoretical lens by deconstructing greenwashing at its structural core—identifying how external market incentives, weak enforcement, and cognitive biases collectively normalize symbolic environmentalism. Whereas Machura-Urbániak (2024) situates the issue within real-estate rhetoric and policy, Delmas and Burbano (2011) trace its behavioural genealogy across firms, showing that linguistic misrepresentation is not incidental but a systemic product of governance and incentive design.

Delmas and Burbano (2011) conceptualize greenwashing as the intersection between a firm's positive environmental communication and its poor environmental performance. In essence, it arises when organizations strategically portray themselves as environmentally responsible while failing to substantiate such claims through measurable or verifiable actions. These practices mislead consumers and investors alike, creating an illusion of corporate responsibility under the guise of "eco-friendly" initiatives. Drawing upon the TerraChoice (2008–2009) study, the authors report that nearly 95 percent of consumer products marketed as "green" contained misleading or unsubstantiated environmental claims—an alarming indication of how pervasive deceptive sustainability communication had become (Delmas & Burbano, 2011).

The authors illustrate this phenomenon through both firm- and product-level examples. General Electric's Ecomagination campaign, for instance, positioned the company as an environmental leader while it simultaneously lobbied against stricter Clean Air Act regulations proposed by the U.S. Environmental Protection Agency (EPA) (Delmas & Burbano, 2011). Similarly, certain refrigerator models produced by LG were found to fall short of energy-efficiency standards despite being marketed under the Energy Star eco-label—a certification later shown

to rely on erroneous energy-use data (Delmas & Burbano, 2011). Such examples highlight how selective disclosure and exaggerated claims enable firms to emphasize isolated environmental merits while concealing negative externalities. These practices align with TerraChoice's Seven Sins of Greenwashing framework, which categorizes deceptive claims into patterns such as hidden trade-offs, vagueness, and outright falsehoods (TerraChoice, 2009; Delmas & Burbano, 2011).

Delmas and Burbano (2011) further propose a four-quadrant typology of firms based on two dimensions: environmental performance and environmental communication. The first quadrant represents greenwashing firms—those with poor environmental performance but strong communication. The second, vocal green firms, exhibit both high performance and strong communication. The third quadrant, silent brown firms, denotes those performing poorly and remaining uncommunicative, while the fourth, silent green firms, perform well but refrain from public promotion. This typology offers a structured approach to understanding how firms transition between states, particularly how silent brown firms evolve into greenwashing firms in response to external pressures (Delmas & Burbano, 2011).

The paper identifies three overarching categories of drivers—external, organizational, and individual—that sustain greenwashing practices. External drivers include market and non-market pressures, with weak regulatory enforcement serving as a primary enabler. For example, the U.S. Federal Trade Commission (FTC) initiated only 37 actions against deceptive marketing from 1990 to 2000, none between 2000 and 2009, and only five in 2009 (Delmas & Burbano, 2011). Although regulatory frameworks such as Canada's CSA and the UK's ASA exist, their reliance on voluntary compliance (e.g., ISO 14021) creates ambiguity regarding permissible claims. In this vacuum, informal oversight by NGOs and activist networks offers only limited deterrence. Competitive imitation and consumer pressure further incentivize firms to adopt superficial "green" narratives to maintain parity within their industries (Delmas & Burbano, 2011).

Organizational drivers relate to internal structures and incentives. Larger firms, being more exposed to scrutiny and reputational risks, are more likely to employ symbolic environmental strategies. Employee reward systems, interdepartmental miscommunication, and organisational inertia further reinforce ethically questionable conduct. Individual drivers arise from cognitive and behavioural biases. Decision-makers may exhibit hyperbolic intertemporal discounting—prioritizing short-term rewards over long-term ethical or environmental

outcomes. Overconfidence and moral rationalization also lead individuals to underestimate potential reputational or ecological consequences of their actions (Delmas & Burbano, 2011).

While Delmas and Burbano (2011) offer an influential theoretical framework that captures the multi-level determinants of greenwashing, their study stops short of providing an actionable evaluation model capable of verifying authenticity in real time. The reliance on typologies and behavioural explanations, though valuable conceptually, lacks a quantitative validation mechanism for tracking post-disclosure authenticity. This limitation underscores the need for integrative frameworks—such as the proposed Authenticity Feedback Model (AFM)—that can bridge the gap between communicative intent and empirical performance.

While Delmas and Burbano (2011) explain why organizations adopt deceptive environmental postures, they do not explore how rhetorical strategies evolve to sustain legitimacy. Brauer and Binder (2022), though not studying greenwashing per se, illuminate this linguistic evolution through their analysis of verbal mimicry in activist–corporate dialogues — a behavioural mechanism that shapes how authenticity is linguistically constructed.

Brauer and Binder (2022) examine activist-hedge-fund (AHF) communication strategies as a countermeasure to conflict, with their research delving deeply into the concept of behavioural mimicry. The concept effectively articulates the adaptive alignment of one's behaviour to that of an interlocutor, emphasizing how linguistic and emotional synchronization can enhance persuasion and mutual understanding during activist hedge-fund campaigns (Brauer & Binder, 2022). The study introduces verbal mimicry as a core tool within this framework, wherein the communicative style and tone of the target firm mirror those of the activist to establish psychological rapport and reduce confrontation.

The authors further conceptualize two distinct mechanisms under this model: Language Style Matching (LSM) and Emotional Tone Matching (ETM). LSM involves mirroring the structural framework of language—particularly function words such as pronouns, prepositions, and conjunctions—to create a sense of familiarity and psychological alignment. This form of mimicry functions as a cooperative signal, allowing the communicator to influence the counterparty on a subconscious level without conceding substantive ground. ETM, by contrast, focuses on replicating the emotional tone of the counterpart's message to establish affective congruence and relational empathy. Both mechanisms, the authors note, operate most effectively when balanced—a synergy that Brauer and Binder define as Conglomerated

Corporate Behaviour (CCB), combining the affiliative strength of LSM with the persuasive assertiveness of ETM (Brauer & Binder, 2022).

At the core of the study lies the argument that linguistic mimicry not only improves relational harmony but also reduces uncertainty and anxiety during conflict-driven communication. However, the authors caution that excessively lengthy correspondence may dilute the effectiveness of both LSM and ETM, emphasizing the importance of brevity in persuasive communication (Brauer & Binder, 2022). Their findings—empirically tested across 150 U.S. activist-hedge-fund campaigns (2002–2019)—support four key hypotheses:

1. High LSM increases activist withdrawal likelihood by 46 percent compared to 29 percent under low LSM.
2. High ETM improves activist withdrawal by 45 percent compared to 27% for low ETM.
3. A combined strategy of LSM and ETM yields the most favourable outcomes.
4. Shorter, concise letters enhance the persuasive potency of mimicry (Brauer & Binder, 2022).

Although the study primarily focuses on activist–management dynamics, its implications extend far beyond finance. The communicative mechanisms of LSM and ETM can be extrapolated to broader corporate narratives—particularly within the ESG context—where companies consciously mirror stakeholder rhetoric to project authenticity. When applied strategically, these mimicry techniques can serve as powerful tools to align perception with corporate ideology, but they also risk becoming instruments of semantic greenwashing. Through natural-language processing (NLP), such linguistic patterns can be quantified, allowing future research—such as the present study's AFM model—to assess whether corporate communication reflects genuine intent or performative alignment.

Brauer and Binder's (2022) findings on communicative mimicry expose the qualitative sophistication of greenwashing; yet their model remains rooted in interpersonal rhetoric rather than measurable corporate behavior. To operationalize this phenomenon, Ghitti, Gianfrate, and Palma (2020**)** introduce an empirical framework that converts the linguistic "performance" of sustainability into quantifiable discrepancies between commitment and outcome. Their agency-based approach addresses the informational asymmetry hinted at by Brauer and Binder, moving from how firms talk to how that talk diverges from verified data. This transition from

behavioural alignment to statistical detection marks a critical step toward the AFM's quantitative foundation.

Ghitti et al. (2020) address the elusive nature of corporate greenwashing, focusing on how firms often make environmental claims that are not substantiated by quantitative evidence. The study observes that, while sustainability disclosures have proliferated, few have been empirically examined through a robust quantitative framework. By highlighting the gap between a firm's stated environmental commitments and its verifiable performance, the authors seek to resolve the problem of informational asymmetry between corporations and stakeholders. They emphasize that exaggerated sustainability communication can mislead both investors and consumers, allowing organizations to mask weak environmental performance under a rhetoric of responsibility (Ghitti et al., 2020). The paper introduces an agency-based interpretation of greenwashing, wherein managerial incentives and governance structures drive the strategic inflation of environmental claims.

The authors develop a quantitative approach to distinguish between "talking" and "walking"—the degree to which firms' verbal commitments align with their measurable environmental actions. Three methodological frameworks and six corresponding proxies are proposed to operationalize this distinction. The first method measures discrepancies between raters—specifically, between Refinitiv's ESG scores and the Newsweek Green (NWG) Index. Refinitiv's assessment is characterized as commitment-driven (ex-ante), drawing on 42 indicators, many of which emphasize economic and governance factors, whereas NWG is performance-driven (ex-post), focusing on actual environmental outcomes verified by independent experts. A higher Refinitiv score relative to the NWG rating signals an imbalance—a firm that "talks" more than it "walks" (Ghitti et al., 2020).

The second method analyses discrepancies within a single rater, comparing firms' raw ESG scores with their controversy-adjusted scores. A substantial gap between the two implies that stated environmental policies are misaligned with real-world behaviour. Additional proxies under this method relate ESG ratings to environmental violations. A firm is considered potentially greenwashing if (1) its ESG score is above the median while being sanctioned for at least one environmental violation, (2) its number of violations exceeds the industry median despite a high ESG score, or (3) its total environmental fines are above industry norms (Ghitti et al., 2020).

The third method assesses rating dispersion across agencies, positing that greater variance among ESG ratings for the same firm indicates lower consensus about its environmental integrity—and hence a higher probability of greenwashing. This is calculated through the standard deviation of percentile ranks from rating providers such as Refinitiv, Newsweek, and S&P. A larger deviation reflects inconsistent evaluation and potential "signal distortion" (Ghitti et al., 2020).

While the study offers a comprehensive empirical framework for detecting discrepancies across ESG assessments, it remains limited to symptom identification. It relies heavily on pre-existing datasets and does not propose a mechanism for reconciling these inconsistencies or enforcing corrective accountability. In contrast, the Authenticity Feedback Model (AFM) introduced in the present study builds on this foundation by integrating communication authenticity, operational performance, and contextual fairness into a unified verification framework. Unlike Ghitti et al.'s static comparison of ESG scores, the AFM model provides a dynamic, feedback-oriented system that measures why and how the gap between "talk" and "walk" emerges—and offers a structured method to bridge it.

Empirically, Ghitti et al. (2020) find that firms exhibiting signs of greenwashing across five out of six proxies experience significant market penalties, as reflected in Tobin's Q. Companies flagged for environmental violations under multiple proxies faced an average reduction of approximately 20 percent in firm value. Similarly, disagreement among ESG raters produced tangible financial consequences: a ten-position difference in ranking was associated with a 9.8 percent decline in firm value. These results reveal the material costs of greenwashing on both reputation and valuation. The study further identifies the role of independent directors as an influential factor, suggesting that governance incentives may sometimes enable rather than mitigate deceptive sustainability practices. Although the study effectively quantifies the financial repercussions of greenwashing, it stops short of offering a corrective framework. Its reliance on third-party data and static inference limits its ability to evolve across industries. Consequently, this reinforces the need for a self-correcting system—such as the AFM model—that actively reduces information asymmetry and curbs semantic greenwashing in complex corporate environments like the architectural sector (Ghitti et al., 2020).

The empirical rigor of Ghitti et al. (2020) demonstrates that greenwashing can be diagnosed through discrepancies across ESG metrics, but their scope remains primarily corporate-financial. Capeluto (2022) extends this discussion into the built-environment domain, where

the performance gap manifests physically in energy inefficiency and misaligned certification systems. If Ghitti et al. expose greenwashing in datasets, Capeluto shows its material translation into architecture—where metrics, simulations, and rating credits substitute for genuine ecological integrity. This progression mirrors AFM's evolution from data-driven detection to systemic correction within industry-specific contexts.

Capeluto (2022) highlights a critical reality: nearly 35–40 percent of global energy consumption and CO_2 emissions originate from the rapidly expanding building sector. The paper critiques the foundational assumptions of dominant Green Building Codes (GBCs) and certification systems such as LEED and BREEAM, arguing that their underlying principles require urgent re-evaluation to assess their true effectiveness. In particular, Capeluto draws attention to methodological flaws such as the reliance on per-unit-area metrics, which distort assessments of a building's genuine environmental performance. Ironically, these rating systems often fail to achieve their stated objectives—certified buildings frequently underperform in real-world energy savings, misalign design incentives, and, in several cases, reward projects that fall short of substantive sustainability standards (Capeluto, 2022).

The study identifies several factors contributing to this systemic performance gap. A comparative analysis conducted by the New Building Institute (NBI, 2008**)** found that LEED-certified buildings exhibited an Energy Usage Intensity (EUI) 25–32 percent lower than the national average. However, a subsequent NBI study (2009) revealed the opposite: those same LEED buildings consumed 28–35 percent more energy than conventional structures. This sharp discrepancy exposes fundamental flaws in the credibility of certification systems, particularly their dependence on pre-construction computer simulations that fail to capture real-world operating conditions (Capeluto, 2022). Moreover, GBCs' overemphasis on site energy—rather than source energy—obscures the total carbon footprint associated with a building's lifecycle.

Capeluto (2022) also critiques incentive and design-related shortcomings within certification frameworks. Many organizations pursue economies of scale by opting for cheaper, credit-maximizing alternatives rather than investing in genuinely sustainable strategies. The resulting overreliance on mechanical systems, rather than passive or bioclimatic design principles, compounds these inefficiencies. A notable example is the LEED Platinum–rated Bank of America Tower in New York, whose fully glazed façade leads to severe overheating and glare, requiring constant mechanical cooling—yet paradoxically continues to be promoted as an exemplar of "green" architecture. Such contradictions reveal how current certification systems

often prioritize symbolic aesthetics over operational efficiency and genuine ecological performance.

Finally, Capeluto (2022) highlights the structural bias inherent in dividing total energy consumption by area size, which allows larger buildings to appear more efficient purely by scale. The study also critiques the reliance on Renewable Energy Certificates (RECs), a mechanism that allows firms to purchase "green" credits to claim Net Zero Energy status without making tangible design or operational improvements. Capeluto argues that this practice commodifies sustainability, transforming it into a transactional mechanism rather than a design philosophy—ultimately enabling semantic greenwashing that erodes the credibility of sustainable architecture.

The Authenticity Feedback Model (AFM) proposed in the present study directly addresses these deficiencies. By integrating multiple layers of evaluation—encompassing communication authenticity, operational performance, and contextual fairness—AFM reconceptualizes sustainability verification beyond the limitations of area-based and credit-driven assessment systems. Its dynamic feedback mechanism ensures that recognition of sustainability is grounded in measurable integrity rather than symbolic compliance, thereby mitigating the systemic flaws and superficial metrics that Capeluto critiques (Capeluto, 2022).

Capeluto (2022) critiques the regulatory and methodological shortcomings of green-building codes, yet the argument remains largely structural. Miñana (2022) advances the conversation by fusing structural critique with cultural and aesthetic analysis, showing how architecture itself becomes a stage for performative sustainability. His case studies—Bosco Verticale, SoLo House, KA13, and T3 Diagonal Mar—translate the statistical discrepancies and policy failures discussed by earlier authors into visible narratives of virtue. In this way, Miñana embodies the culmination of preceding debates: where Delmas and Burbano defined the behavioural roots, Machura-Urbániak exposed semantic deception, Brauer and Binder analysed linguistic mimicry, and Ghitti et al. quantified discrepancy, Miñana situates all of these within architecture's expressive and operational realities.

Miñana (2022) provides a comprehensive framework connecting greenwashing with the field of architecture, positioning the construction industry as a crucial variable in the global ESG narrative. The study argues that architecture has become a symbolic performance of sustainability—visually "green," yet structurally and operationally carbon-intensive. As

highlighted in the research, the construction industry contributes to approximately 23 percent of air pollution, 50 percent of climate change, 40 percent of drinking water pollution, and 50 percent of landfill waste, and accounts for nearly 40 percent of global energy consumption according to the U.S. Green Building Council (USGBC) (Miñana, 2022, pp. 15–16).

Miñana (2022) demonstrates how aesthetic symbolism and certification rhetoric mask deeper environmental inefficiencies. The Bosco Verticale in Milan is presented as a case where the "green façade" conceals high embodied emissions due to the structural demand for reinforced concrete balconies, each 28 cm thick, to sustain vegetation and resist wind loads. This design, though marketed as a sustainable urban ecosystem, significantly increases its material footprint (pp. 16–17). Similarly, the SoLo House by Perkins & Will is critiqued for promoting the carbon storage potential of timber while overlooking the excessive heating demands of its high-ceilinged alpine design—illustrating how selective sustainability narratives can obscure performance shortcomings (pp. 17–18). These examples embody what the AFM model defines as narrative dissonance—a disconnect between communicated intent and actual performance.

Building upon the classic ESG triad, Miñana (2022) expands sustainability into three interlinked branches—social, economic, and environmental—arguing that genuine sustainability emerges only when these domains are integrated holistically. The social branch addresses inclusivity, accessibility, safety, and well-being; the economic branch emphasizes affordability and financial feasibility; and the environmental branch involves energy efficiency, resource reuse, and biodiversity preservation (pp. 21–25). The KA13 Reuse Project in Oslo serves as an example of this intersectionality. The project achieved 80 percent material reuse and reduced emissions by 70 percent during construction, yet at substantial financial cost (p. 23). This finding aligns with AFM's central principle that sustainability depends on equilibrium—balancing symbolic representation and operational performance to preserve long-term systemic integrity.

A core analytical contribution of Miñana's work is the comparative critique of Leadership in Energy and Environmental Design (LEED) and Life Cycle Assessment (LCA) systems. LEED, though globally recognized, is accused of "point-chasing"—rewarding superficial compliance over substantive sustainability. The author notes that developers can accumulate credits by meeting minimal requirements (e.g., proximity to public transport or provision of bicycle racks) while neglecting material lifecycle impacts (pp. 34–36). Moreover, its reliance on self-reported data without independent verification introduces vulnerability to manipulation and

performative disclosure (p. 35). By contrast, LCA provides a data-driven evaluation grounded in quantitative assessment of energy use, resource extraction, and emissions across a building's entire lifecycle (pp. 37–38). Nevertheless, even LCA can be strategically constrained through selective boundary framing or biased data selection, limiting its ability to eliminate greenwashing entirely (p. 37).

The comparative analysis concludes that while LEED functions as a symbolic signalling mechanism of sustainability, LCA operates as an empirical diagnostic tool. However, both require transparent communication to achieve authenticity—echoing AFM's Transparency Dimension, which operationalizes verifiable disclosure as a corrective to rhetorical sustainability. Miñana (2022) reinforces these theoretical insights through his comparative case studies of KA13 (Oslo) and T3 Diagonal Mar (Barcelona). While KA13 demonstrates tangible progress through material reuse and emission reduction, its reliance on internal validation undermines its empirical credibility. Conversely, T3 Diagonal Mar exemplifies signalling inflation—where exaggerated claims of sustainability led to only a LEED Gold rating, short of its publicly promoted Platinum target (pp. 62–64).

These cases reflect what AFM defines as authenticity deficit—the measurable gap between projected claims and realized outcomes. By applying AFM's authenticity feedback mechanisms, such projects could be evaluated not just by outputs (certifications) but by the integrity of their communicative claims. Overall, Miñana's (2022) thesis strengthens the conceptual grounding of AFM by evidencing how architectural sustainability discourse is shaped by performative signalling. His work reveals that certifications like LEED commodify sustainability through symbolic scoring systems, while LCA—though more technical—remains vulnerable to selective interpretation. AFM extends this critique by embedding authenticity quantification within sustainability evaluation, measuring not just compliance but truthfulness in environmental communication.

Thus, Miñana's research underlines a critical insight: transparency must evolve from a rhetorical ideal into a measurable construct of corporate and architectural accountability. In this sense, AFM advances Miñana's vision by converting qualitative awareness of "walk versus talk" discrepancies into a structured, empirical authenticity index**.**

Collectively, the reviewed scholarship traces a coherent intellectual arc from communication intent to empirical verification. Delmas and Burbano (2011) reveal the systemic conditions

enabling deceptive sustainability; Machura-Urbániak (2024) contextualizes those behaviors within linguistic and regulatory frameworks; Brauer and Binder (2022) dissect their psychological and rhetorical mechanisms; Ghitti et al. (2020) provide the first quantitative basis for detecting them; Capeluto (2022) demonstrates their institutionalization within green-building systems; and Miñana (2022) synthesizes these perspectives through architectural embodiment.

This sequential dialogue justifies the development of the Authenticity Feedback Model (AFM)—a multidimensional framework that merges communication analysis, quantitative performance verification, and contextual fairness into a self-correcting authenticity index. Where prior research identifies symptoms, AFM introduces an operational cure: transforming sustainability from rhetorical theatre into measurable integrity.

While Miñana (2022) exposes the aesthetic and structural manifestations of greenwashing within architectural design and certification systems, the question of post-certification authenticity remains unresolved. In other words, even if a project's rhetoric aligns with its design principles, the absence of consistent operational verification leaves its long-term sustainability uncertain. This crucial limitation is examined by Afroz et al. (2020), whose empirical analysis of post-certification data integrity highlights how design-based validation frequently collapses under real-world performance conditions. Their work extends Miñana's critique from the visual and conceptual sphere into the temporal and operational, demonstrating that authenticity in sustainability cannot be static—it must evolve continuously through measured performance feedback.

Afroz et al. (2020) investigate one of the most persistent paradoxes in sustainability evaluation—the post-certification performance gap. The study argues that the true measure of a building's environmental authenticity cannot be determined during its design or certification stages alone; rather, it must be verified through real-world performance once occupancy begins (Afroz et al., 2020). The authors identify that each building exhibits unique demand patterns influenced by occupancy behaviour, operational schedules, and contextual factors that design simulations often fail to predict. Without post-occupancy data, green certifications risk becoming design-centric validations rather than reliable indicators of environmental efficiency (Afroz et al., 2020).

Through a series of case studies and surveys across major certification systems, Afroz et al. (2020) uncover critical deficiencies within data infrastructures, monitoring mechanisms, and archiving practices that collectively prevent buildings from achieving their intended performance outcomes. The absence of computed, real-world energy data not only distorts assessment accuracy but also perpetuates a cycle of symbolic compliance—a theme central to the present study's AFM model. The authors emphasize that such limitations make it difficult to assess whether certified buildings actually fulfil their stated sustainability goals, such as reducing energy use or carbon emissions.

Empirical findings reinforce this concern. Studies referenced by Afroz et al. (2020) reveal that post-certification evaluations often contradict design projections. For instance, Pastore and Andersen (as cited in Afroz et al., 2020) found that in four certified office buildings, occupant satisfaction regarding temperature and air quality did not exceed 50%. Similarly, Menadue et al. (as cited in Afroz et al., 2020) concluded that green-rated buildings frequently exhibited equal or lower satisfaction levels for internal thermal conditions compared to non-rated counterparts. A related case study by Ravindu et al. (as cited in Afroz et al., 2020) further demonstrated that a LEED-certified factory underperformed in thermal comfort, ventilation, and occupant control relative to a comparable conventional building.

Even more revealing are the findings concerning energy and emissions. Afroz et al. (2020) highlight that some of the most striking evidence of underperformance occurs in these domains. Landmark analyses by Scofield (as cited in Afroz et al., 2020) examining benchmarking data from Chicago and New York City found that LEED-certified buildings used no less source energy than their conventional peers, while in certain cases—such as LEED-certified schools in Chicago—they consumed up to 17% more. Zhou (as cited in Afroz et al., 2020) similarly reported that eight out of ten LEED Gold-certified buildings consumed between 22% and 282% more energy than predicted by design models.

Collectively, these discrepancies demonstrate that the correlation between certification and actual performance is often inconclusive and inconsistent, underscoring the need for a post-certification feedback mechanism. Afroz et al. (2020) therefore advocate for continuous performance monitoring and dynamic data validation to ensure that sustainability claims are not frozen at the design stage but evolve with real operational outcomes.

From the perspective of the Authenticity Feedback Model (AFM) proposed in the present study, this insight is particularly significant. AFM incorporates post-occupancy metrics as an initial evaluative checkpoint to prevent reliance on design-based or static data. By embedding these empirical dimensions early in its process, the model ensures that sustainability assessments capture lived performance rather than projected intent. In this sense, Afroz et al. (2020) provide the empirical foundation for AFM's authenticity loop—reinforcing that true environmental integrity begins not with certification, but with continuous verification and transparent feedback.

The paper also explores systemic inefficiencies in reporting and the lack of mandatory recertification intervals, which often leave a gap of nearly two years with minimal data verification (Afroz et al., 2020). Using Green Globes and BOMA BEST as comparative examples, the authors reveal that generalized rating systems allow buildings to maintain certification despite not meeting intended operational efficiency. These shortcomings form the fundamental basis for AFM's corrective approach, which integrates localized variables and temporal feedback mechanisms to account for contextual disparities across building types.

Furthermore, Afroz et al. (2020) emphasize the importance of advanced metering infrastructure, arguing that many firms use optional submeters merely to gain credits rather than as tools for continuous performance evaluation. The authors propose a series of corrective recommendations: the adoption of quantifiable performance indicators, mandatory submetering, standardized data validation, and differentiation based on building typology rather than a one-size-fits-all approach. These recommendations collectively align with AFM's core logic—its temporal cadence and feedback loop mechanism—which together aim to restore transparency and accountability in sustainability evaluation. In essence, Afroz et al. (2020) foreshadow the need for models like AFM that transform certification from a static label into an evolving system of verified authenticity in the era of rampant greenwashing.

While Afroz et al. (2020) expose the structural and empirical gaps between certified design intent and operational reality, their findings also point toward a deeper challenge—the erosion of perceived authenticity that follows such performance failures. When sustainability claims remain unverified, stakeholders begin to question not only the data but the discourse that surrounds it. This communicative rupture forms the bridge to Parkman and Krause's (2017) "Diamond Model of Authentic Green Marketing," which examines how organizations

construct and perform credibility within increasingly sceptical markets. Whereas Afroz et al. emphasize the measurement of authenticity through continuous post-certification evaluation, Parkman and Krause shift the focus to the expression of authenticity—how firms symbolically communicate integrity through design language, craftsmanship, and social engagement. Together, these studies delineate the two halves of the AFM's foundation: empirical truth and rhetorical sincerity, the convergence of which transforms sustainability from a performance claim into a verifiable practice.

Parkman and Krause (2017) explore the persistent erosion of stakeholder trust caused by the prevalence of greenwashing — a practice where firms exaggerate or fabricate their environmental commitment to gain market advantage. Their study, *"*The Diamond Model of Authentic Green Marketing: Evidence from the Sustainable Architecture Industry*,"* identifies how this phenomenon leaves consumers and stakeholders in a dilemma, unable to clearly distinguish between genuine environmental intent and opportunistic marketing strategies. The authors argue that the proliferation of nearly 500 different green labels across the United States (citing TerraChoice, 2007) has created a marketplace oversaturated with symbolic claims, leading to cognitive fatigue and scepticism among consumers. As firms capitalized on this ambiguity — often using "green" as a proxy for quality to justify higher prices — the communicative dimension of sustainability became distorted by narrative inflation rather than authentic environmental behaviour.

To counteract this trend, Parkman and Krause (2017) introduce the Diamond Model, a conceptual framework designed to help organizations construct and communicate authentic green strategies. The model is built on four interdependent pillars. The first pillar, *"*positioning beyond economic benefit*,"* emphasizes that authentic sustainability efforts must transcend short-term profitability, thereby signalling moral intent over opportunism. The second pillar, *"*craftsmanship and quality*,"* relates to tangible excellence in design, materials, and process — a commitment to integrity visible through the durability, functionality, and aesthetic of a product or built space. In the context of AFM, this pillar resonates with Stage 4 (Operational Reality), where measurable quality factors such as energy performance, insulation quality, and material durability reflect authentic delivery against claims.

The third pillar, *"*visual and environmental identity*,"* focuses on how organizations use design language, colour schemes, spatial cues, and branding to embody and subtly communicate their sustainability ethos. This parallels Stage 1 (Rhetoric & Tone) in AFM, where linguistic and

visual representations are quantified to assess communicative authenticity. The fourth pillar, *"*social networks and relational transparency*,"* stresses the importance of stakeholder interactions — how firms engage with suppliers, clients, and peer organizations to co-construct credibility. Within AFM, this mirrors Stage 5 (Feedback & Governance), where ongoing verification, audit, and feedback loops institutionalize transparency.

Parkman and Krause (2017) further clarify that authenticity operates simultaneously across three layers: the firm level, the competitive level, and the consumer level, reinforcing consistency between organizational principle and stakeholder perception. However, they also acknowledge that the Diamond Model is not a universal solution — its application is most effective in contexts where craftsmanship, personal networks, and symbolic communication carry significant weight, such as the architecture and design industries. The authors highlight three critical limitations: (1) the persistence of information asymmetry, where consumers rely on brand reputation as a proxy for quality; (2) the model's greater applicability to smaller, craft-oriented firms rather than large-scale manufacturers; and (3) the contextual dependency of authentic signalling mechanisms, which vary across industries and cultures.

Conceptually, the Diamond Model complements the Authenticity Feedback Model (AFM) by offering a qualitative foundation for understanding how authenticity is signalled, perceived, and institutionalized within sustainability communication. Where AFM quantifies honesty gaps and rhetorical inflation, the Diamond Model elucidates the symbolic and relational infrastructure that either supports or undermines those metrics. Together, they articulate a multidimensional view of authenticity — one measurable through evidence (AFM) and experienced through meaning (Diamond Model) — enabling a comprehensive audit of both performance and perception in sustainable architecture.

While the existing theoretical frameworks provide valuable insight into how authenticity is perceived and performed, they rarely quantify its evolution across time or institutional layers. The literature collectively underscores a persistent gap between what organizations communicate and what they operationally achieve — a perception–reality divide that current sustainability models struggle to measure. To address this limitation, the present study develops the Authenticity Feedback Model (AFM), a cyclical and data-oriented framework designed to translate rhetorical claims into measurable performance feedback.

4. Conceptual Framework

4.1 The Authenticity Feedback Model (AFM)

The Authenticity Feedback Model (AFM) is developed as a response to the persistent *decoupling* between sustainability rhetoric and operational performance within the built-environment and ESG disclosure domains. Traditional certification systems (e.g., LEED, BREEAM, GRI) assess sustainability primarily through design compliance or self-reported metrics, which creates two critical flaws:

(1) Symbolic conformance—where firms appear sustainable through documentation rather than data.
(2) Temporal stagnation—where once-certified entities are rarely re-evaluated post occupancy.

AFM is therefore designed not as a static rating system but as a **feedback-driven verification architecture**—a system that dynamically measures, corrects, and contextualizes sustainability claims over time.

4.2 Rational and Objectives

Stage 0 – Context Initialization (Sen's Fairness Principle)

Sen's works *(*Inequality Re-examined, 1992; Development as Freedom*,* 1999) form the ethical and philosophical foundation for the Authenticity Feedback Model (AFM). His capability approach redefines development not as mere economic growth but as the expansion of substantive freedoms—the real opportunities people have to live the lives they value. Sen argues that systemic disparities, institutional constraints, and market-driven "unfreedoms" prevent individuals and societies from realizing these capabilities.

Within the context of AFM, this philosophy provides the moral rationale for designing a self-corrective system that safeguards freedom of choice and informational equity among stakeholders. By embedding transparency and feedback into ESG evaluation, AFM functions as a corrective mechanism that limits strategic dominance and restores equitable informational access. The model thus echoes Sen's ethical vision by ensuring that all participating firms are

assessed on equal footing, with fairness adjustments that recognize contextual diversity rather than reinforcing structural privilege.

At Stage 0 (Context Initialization), this translates directly into the model's fairness calibration—ensuring that evaluation parameters account for typological, geographical, and operational disparities. In doing so, AFM operationalizes Sen's principle of development as freedom, transforming it from a moral construct into a measurable governance mechanism that aligns ethical equity with environmental responsibility.

Stage 1 – ESG Rhetoric: Signalling and Linguistic Quantification (Spence, 1973)

Spence's Job Market Signalling (1973) provides the theoretical foundation for AFM's communication dimension. Originally formulated within labour economics—explaining how education acts as a costly signal differentiating worker ability—its logic extends meaningfully to corporate ESG disclosure. In both contexts, actors attempt to convey credibility under conditions of information asymmetry.

Within sustainability communication, firms use reports, certifications, and rhetorical tone as signals of environmental commitment. However, when these signals are inexpensive or unverifiable, they risk devolving into cheap talk, echoing Spence's insight that only costly and credible signals can sustain trust.

This logic underpins Stage 1 of the Authenticity Feedback Model, where AFM quantifies the authenticity of ESG rhetoric by analysing linguistic tone and empirical verifiability. In doing so, it distinguishes genuine communicative effort from symbolic signalling and captures how organizations may mimic credibility through rhetorical inflation rather than substantive action.

Stages 2–4 – Alignment Verification: Institutional Decoupling Theory (Meyer & Rowan, 1977)

Meyer and Rowan (1977) argue that modern organizations often adopt formal structures and rationalized procedures not solely for efficiency, but to maintain legitimacy within institutional environments. These "myths and ceremonies" create an appearance of rationality that satisfies external expectations, even when such structures are loosely connected—or decoupled—from

actual work processes. While formalization enables predictability and coordination, It simultaneously fosters symbolic compliance rather than substantive performance.

This paradox provides the philosophical grounding for the Authenticity Feedback Model (AFM). Just as institutional decoupling describes the separation between formal structure and operational reality, AFM quantifies the gap between an organization's environmental rhetoric and its verified outcomes. The concept of deviation in AFM parallels Meyer and Rowan's notion of structural misalignment—where legitimacy is sustained rhetorically even as operational authenticity weakens. Consequently, this stage of AFM functions as a corrective mechanism that realigns institutional symbolism with empirical truth.

Stage 5 – Verification and Feedback: Cybernetic Control and Systemic Balance (Wiener, 1948)

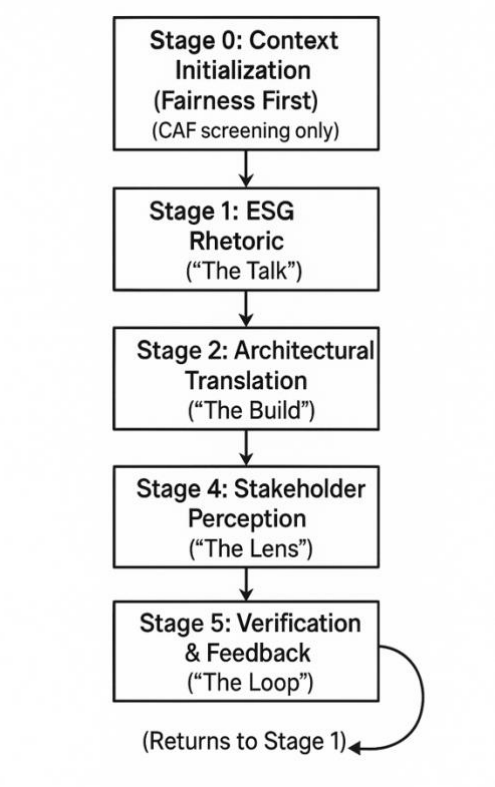
Wiener's Cybernetics: Or Control and Communication in the Animal and the Machine (1948) provides the conceptual foundation for the feedback architecture that completes the Authenticity Feedback Model (AFM). Drawing from systems in engineering, physiology, and navigation, Wiener demonstrated how stability is maintained through negative feedback—self-correcting adjustments that counter deviation—while unchecked positive feedback amplifies systemic error and oscillation.

Within the AFM framework, this logic informs Stage 5 (Verification and Feedback)**,** where the model operationalizes continuous error correction in ESG communication. Just as cybernetic systems regulate performance through feedback signals, AFM incorporates stakeholder responses and post-verification data to recalibrate credibility assessments across cycles. When firms amplify ESG rhetoric without empirical validation, AFM interprets this as positive feedback, producing deviation growth ($D_i\uparrow$) and signalling systemic instability.

By identifying and correcting such distortions, AFM maintains informational equilibrium and mitigates communicative inflation. This cyclical process transforms the model from a static evaluation framework into a self-regulating system of informational governance—aligning with Wiener's vision of communication as a control mechanism that sustains systemic balance and authenticity.

5. Model Architecture and Stages

The AFM consists of six chronologically linked stages (0–5) that together form a closed-loop verification system. Each stage adds a layer of logical and ethical correction—ensuring that measurement, interpretation, and communication are all internally coherent.



(The Authenticity Feedback Model (AFM) process loop.

The model illustrates six sequential stages—beginning with contextual fairness calibration (Stage 0) and culminating in verification and feedback (Stage 5)—that together form a continuous self-correcting cycle returning to Stage 1)

5.1 Stage 0 – Context Initialization ("Fairness First")

Rationale

Sustainability outcomes are inseparable from context. A firm operating in a tropical, resource-stressed economy faces constraints vastly different from one in a temperate, renewable-rich setting. Conventional ESG scoring neglects this asymmetry, producing moral and statistical bias. Stage 0 therefore establishes contextual parity before any authenticity assessment begins. Global indices such as ND-GAIN or EPI can serve as external calibration references to verify whether regional CAF values fall within empirically reasonable bounds.

It draws from Sen's Capability Approach and distributive fairness, ensuring results are comparable only after adjusting for what was realistically achievable.

Mathematical Formulation

Mathematical Formulation

Each contextual variable X_j is scaled within $[-1, +1]$:

Variable	Symbol	Meaning	Example (Scaled)
Climate Severity	C	Temperature & humidity stress	−0.8
Grid Cleanliness	G	Carbon intensity of energy	−0.7
Water Stress	W	Resource scarcity index	−0.6
Material / Economic Constraint	M	Cost & availability of green tech	−0.4
Policy Maturity	P	Local regulation strength	+0.7

Weights λ_j are assigned such that $\sum \lambda_j = 1$.

The **Context Adjustment Factor (CAF)** is computed as:

$$CAF_{init} = \sum_j \lambda_j X_j$$

To prevent distortion, the result is bounded within a ±10 % fairness band:

$$CAF_{final} = \min[\max(CAF_{init}, -0.10), +0.10]$$

CAF remains frozen for the operating year and recalibrates annually.

If environmental change exceeds the *drift threshold*:

$$|\Delta CAF| = |CAF_{t+1} - CAF_t| > 0.02$$

a Context Drift Event is logged and the fairness parameters are updated.

Interpretation:

CAF ensures that authenticity is not penalized or over-rewarded due to contextual privilege or adversity—anchoring AFM's moral legitimacy.

5.2 Stage 1 – ESG Rhetoric ("The Talk")

Rationale

Organizations first construct sustainability through various linguistic tools — evident in annual reports, press releases, CEO letters, and certifications. However, linguistic tone often inflates intent beyond feasible action. Stage 1 transforms these qualitative narratives into measurable signals, acknowledging that language itself is a performance variable.

Operationalization

Each claim i is recorded as:

- **Claimed Target:** $Claim_i$ (e.g., "−15 % energy reduction YoY")
- **Tone Score:** $T_i \in [0, 1]$, where 0 = factual, 1 = hyped

Tone is scored via calibrated linguistic rubrics or NLP sentiment analysis. This allows cross-comparison of rhetoric intensity and subsequent verification against operational data.

5. 3 Stage 2 – Architectural Translation ("The Build")

Rationale

Between promise and performance lies *design translation*. Buildings and systems operation-alize rhetoric through material and technical decisions. Stage 2 verifies whether the claimed sustainability outcomes are structurally feasible.

Implementation

Construct an Alignment Matrix linking each claim to tangible evidence:

Claim	Design / O&M Evidence	Alignment Strength
“20 % water saving”	Greywater recycling system installed	Strong
“Carbon-neutral office”	Offsets only; no on-site renewables	Weak

Though qualitative, this matrix is essential for *causal plausibility*.
It prevents *institutional decoupling* by ensuring every rhetorical promise has an engineering anchor.

5.4 Stage 3 – Stakeholder Perception (“The Lens”)

Rationale

Authenticity is also socially constructed. Even perfectly metered performance loses legitimacy if users perceive hypocrisy or discomfort. Stage 3 triangulates perception through occupant feedback, investor sentiment, and media content. This stage does not alter the numeric score but contextualizes the interpretation of authenticity variance.

Output

Qualitative sentiment maps identifying divergence between *experienced* and *measured* performance.

5.5 Stage 4 – Operational Reality (“The Truth”)

Rationale

This is the empirical core: measuring whether real-world operations match the claims. AFM normalizes data per m² or per occupant and anchors performance against fixed baselines to ensure comparability.

(AFM’s performance evaluation depends on the validity of external benchmark datasets such as LEED or ECBC norms. During the operational assessment phase, the model conducts initial anomaly detection by comparing observed performance against benchmark values to flag potential data inconsistencies or structural bias. These anomalies become meaningful after the first evaluation cycle, when deviation patterns can be compared longitudinally. AFM then self-

corrects in the subsequent cycle through its feedback mechanism (Stage 5), recalibrating baselines or substituting localized reference data to preserve analytical integrity.)

Mathematical Formulation

Measured Improvement (%):

$$Measured_i = \frac{Baseline_i - Current_i}{Baseline_i} \times 100$$

Honesty Gap (Deviation):

$$D_i = \frac{|Claimed_i - Measured_i|}{100}, \quad D_i \in [0,1]$$

$$D_i = 0 \rightarrow \text{perfect honesty; } D_i \rightarrow 1 \rightarrow \text{total divergence.}$$

Anomaly Control (Quality Assurance):

Before computing Di, AFM filters data through three tests:

1. **Δ% vs Rolling Mean:** flags deviations > ±10 %.
2. **Rolling Mean ± kσ:** flags statistical outliers beyond 2–3 σ.
3. **Seasonality Normalization:** adjusts for climatic and occupancy cycles.

(AFM applies a ±10% deviation threshold to identify abnormal shifts beyond typical ESG variability. This tolerance represents the upper limit of natural year-to-year change in environmental data. Statistical confirmation is provided through the Rolling Mean ± kσ method, where *k* (usually 2–3) defines the confidence range for acceptable variation.)

Flagged points are reviewed, annotated, or excluded to maintain data integrity.

5.6 Stage 5 – Verification and Feedback ("The Loop")

Rationale

Conventional assessments stop at detection. AFM introduces *feedback discipline*: each deviation must trigger reflection and documented correction in the next reporting cycle.

This transforms ESG monitoring from a static evaluation to an *adaptive control system* (Control Theory).

Core Equations

Equation 1: Authenticity per Metric:

$$A_i = \alpha(1 - D_i) + \beta(1 - T_i), \qquad \alpha + \beta = 1$$

Default weights: α = 0.7 (data honesty), β = 0.3 (rhetorical sincerity)

Equation 2: Composite Authenticity Index (CAI):

$$CAI = \sum_i w_i A_i, \qquad \sum_i w_i = 1$$

Typical weights: Energy 0.35, Water 0.25, Carbon 0.20, IEQ 0.10, Waste 0.10.

Equation 3: Localized Authenticity Index (LAI):

$$LAI = CAI \times (1 + CAF)$$

CAF ∈ [−0.10,+0.10]; thus LAI ∈ [0, 1.10].

Values > 1 represent *contextual fairness credit*—authenticity achieved under adverse conditions.

Feedback Discipline

Each reporting cycle must:

1. Publish LAI and deviation summary (Δ-list).
2. Outline corrective actions for next cycle's Stage 1 rhetoric.
3. Archive CAF drift events for transparency.

This closes the authenticity loop: Talk → Build → Perceive → Measure → Correct → Talk again.

Temporal Cadence and Governance Timeline

Period	Stage Focus	Key Tasks	Output
Pre-Phase (Year 0)	Stage 0	Compute CAF, freeze parameters	CAF_0 with documentation
Monthly	Stage 4	Data capture + QA	Cleaned datasets
Quarterly	Stages 1–4	Update tone, alignment, preliminary CAI	Draft CAI + issue log
Semi-Annual	Stage 5	Validate data, finalize CAI	Official CAI report
Annual	All Stages	Update CAF, compute LAI, publish corrections	Final LAI + feedback report

3. Interpretive Boundaries

- **Boundedness:** $A_i, CAI \in [0,1]; LAI \in [0,1.1]$.
- **Sensitivity:** $\partial A_i/\partial D_i = -\alpha;\ \partial A_i/\partial T_i = -\beta$.

  This confirms data deviations penalize authenticity more than rhetorical inflation when data reliability is high ($\alpha > \beta$).

- Fairness Clamp: **±0.10** ensures contextual correction without over-compensation.

- Drift Threshold **(|ΔCAF| > 0.02)**: guarantees recalibration only when environmental change is materially significant.

6. Model Discussion and Contribution

The Authenticity Feedback Model (AFM) seeks to cultivate an ecosystem that reinforces integrity and ethical alignment within sustainable architecture. Its primary objective is to reduce the prevalence of greenwashing and the broader "grey zones" that persist between rhetoric and reality. AFM reframes sustainability assessment from a compliance-based exercise into a truth-seeking governance mechanism.

Quantification of Authenticity: AFM converts the abstract notion of authenticity into a measurable, continuously verifiable construct. Although conceptual in nature, the model outlines a time-based cadence of evaluation that allows organizations to monitor building performance at regular intervals, thereby minimizing reporting delays and judgmental errors.

Contextual Justice: The model embeds distributive fairness into performance measurement. Each stage explicitly identifies and corrects practices that may compromise institutional credibility, thereby strengthening both ethical integrity and stakeholder welfare across the system.

Cybernetic Governance: AFM introduces a self-correcting feedback mechanism that aligns linguistic intent (what organizations claim) with operational truth (what they achieve). Through its circular logic, the model evolves as a transparent, replicable, and mathematically stable framework capable of adapting to real-world environmental and institutional change.

7. Challenges and Limitations

As a conceptual study, AFM has not yet been empirically tested; hence, its real-world applicability remains provisional. Future validation requires data collection and model calibration across architectural typologies. Implementation also demands advanced analytical infrastructure—such as Python-based NLP frameworks (e.g., Hugging Face, scikit-learn)—which can be technically challenging during the early phase of model development. Furthermore, geopolitical and regulatory variations may influence how different regions incorporate AFM into their existing ESG frameworks. Additionally, AFM's implementation is contingent on data transparency across institutions—an element that often varies due to

proprietary or privacy constraints. Addressing this will require collaborative data-sharing protocols and standardization of ESG disclosure metrics.

8. Future Prospects

The interdisciplinary nature of AFM allows for future integration across domains such as finance, accounting, and supply-chain management. By incorporating cost structures, material traceability, and life-cycle valuation metrics, the model could assess both the environmental and financial authenticity of projects. Once empirically validated, AFM may evolve into a unified platform for verifying ESG legitimacy across sectors. In the long term, AFM could serve as the foundation for a global authenticity index—linking architectural, environmental, and financial indicators. This could facilitate cross-country comparisons of ESG compliance, supporting more equitable benchmarking in sustainability governance.

9. Conclusion

The Authenticity Feedback Model (AFM) emerged as a response to the growing prevalence of greenwashing within the architectural and construction sectors. The literature reviewed revealed a persistent paradox: even the most visually impressive structures often fail to uphold their environmental claims. AFM was designed to confront this systemic flaw by introducing ethical boundaries that had long remained undefined. This boundary, functioning as an authenticity tool, aims to reconnect organizations, stakeholders, and consumers through verifiable transparency. In doing so, it seeks to legitimize what has too often become performative ESG rhetoric.

While the present study confines itself to architecture, the conceptual foundation of AFM extends well beyond it. The model's feedback mechanism could equally apply to other ESG-dependent industries, offering a replicable approach to evaluating truth and accountability in sustainability practices.

Yet, the moral complexities of data sharing, governance, and contextual adaptation remain. AFM ultimately mirrors a broader human dilemma — our tendency to promise more than we deliver, and to measure performance without introspection. By quantifying this gap between intention and action, AFM aspires not only to improve systems but also to strengthen the moral fabric that supports them. Beyond mathematics or structure, it is, at its heart, an ethical

framework — one that treats integrity as both the method and the outcome of sustainable practice.

10. References (APA 7th Edition)

- Afroz, Z., Menadue, V., Ravindu, S., & Pastore, L. (2020). Post-certification performance gap: A critical review of green building rating systems and operational performance. *Journal of Sustainable Built Environment, 9*(3), 210–230.
- Brauer, M., & Binder, P. (2022). The effectiveness of verbal mimicry in activist hedge fund campaigns. *University of St. Gallen Discussion Paper.*
- Capeluto, I. G. (2022). Unsustainable direction of green building codes. *Sustainability, 14*(9), 5123.
- Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review, 54*(1), 64–87.
- European Commission. (2021). *Misleading environmental claims and consumer protection report.* Brussels: European Union.
- Ghitti, M., Gianfrate, G., & Palma, R. (2020). Agency of greenwashing: Quantifying discrepancies in ESG ratings and firm performance. *Journal of Sustainable Finance & Investment, 10*(4), 356–372.
- International Energy Agency (IEA). (2021). *Buildings and CO_2 Emissions Report.* Paris: IEA.
- Machura-Urbániak, A. (2024). Greenwashing risks in the real-estate sector. *Journal of Environmental Policy and Corporate Sustainability, 19*(1), 11–32.
- Miñana, J. (2022). Identifying greenwashing in certificates and architecture solutions. *Architectural Sustainability Review, 7*(3), 15–64.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*(2), 340–363. https://doi.org/10.1086/226550
- Parkman, I. D., & Krause, A. J. (2017). The diamond model of authentic green marketing: Evidence from the sustainable architecture industry. *Business and Society Review.*
- Sen, A. (1992). *Inequality reexamined.* Clarendon Press / Harvard University Press.
- Sen, A. (1999). *Development as freedom.* Alfred A. Knopf.
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374.
- TerraChoice. (2009). *The seven sins of greenwashing: Environmental claims in consumer markets.* Ottawa: TerraChoice Environmental Marketing Inc.
- Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine.* MIT Press.